Synthesis of an Artificial Homo-C Nucleotide and Investigation of Its Base-Pairing Ability

Dissertation

zur

Erlangung der wissenschaftlichen Doktorwürde

(Dr. sc. nat.)

vorgelegt der

Mathematisch-naturwissenschaftlichen Fakultät

der

Universität Zürich

von

Benno K. Bischof

von

Eggersriet/ Grub SG

Promotionskomitee

Prof. Dr. Jay S. Siegel (Vorsitz)

Prof. Dr. John A. Robinson

Prof. Dr. Oliver Zerbe

Zürich, 2009

Contents

Summary

Questions about the emergence of life has fascinated humankind for all times. On the basis of *Darwin's* generally accepted theory of evolution, countless hypotheses try to explain the evolutionary progression. Especially the *Origin of Life*, that is the formation and characteristics of the first "living" unit, sparks our interest.

In this work, I attempted to synthesize a possible precursor to the present-day hereditary material (DNA) and to investigate its chemical and physicochemical properties. Today's DNA consists of four different nucleotide building blocks, which are in turn subclassified into two categories: purines and pyrimidines. A hypothesis, posited by *Siegel* and *Tor*, proposes a system based on four pyrimidine bases whereof the present-day nucleic acids could have arisen. One single building block could have been the source of these four suggested nucleobases, whereas the two contemporary pyrimidine bases C and U plus two yet unknown nucleotides D and E would pair in a *Watson/Crick* -mode. D and E exhibit two exocyclic amino nucleotides (EANs) possessing an exocyclic amino-glycosidic junction between the pyrimidine and the sugar moiety (see *Figure 1*). As harbingers of A and G, respectively, D and E should show similar chemical properties concerning strength of base paring, as well as selectivity in binding of complementary bases.

In order to investigate and confirm these prerequisites, a model compound of nucleotide D was synthesized and incorporated into oligonucleotides. Due to problems of synthesis and stability, a homo-C analogue was chosen as a model compound. Compared to EANs, a homo-C nucleotide features a

more stable methylene-bridged glycosidic bond, but should be sterically and electronically quite similar.

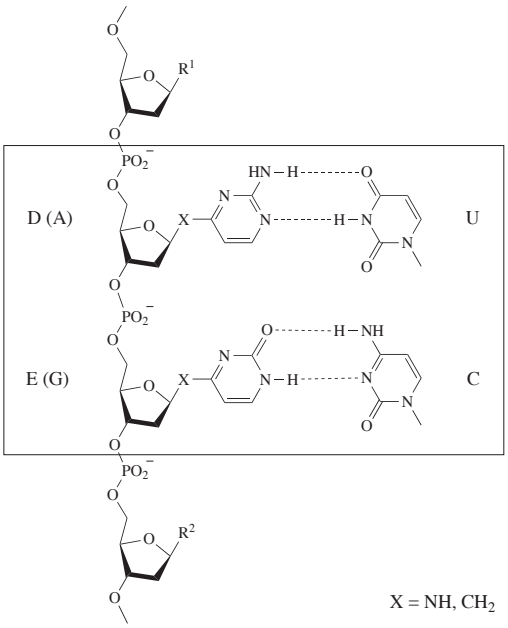


Figure 1: Proposal of a DNA precursor.

I have been able to determine the base pairing properties of homo-*C* nucleotide D qualitatively by DNA melting experiments of different complementary, double stranded oligonucleotides. The examinations demonstrated that D has characteristics similar to its analog A and is able to bind U in *Watson/Crick* -mode; however, the binding ability of D seems to be highly dependent on its proximity. Presumably, strong dipole inducing flanking bases are necessary to stabilize the flexible homo-*C* nucleotide D in a stacked conformation. If this more or less stable stacking is assured, an effective binding can be observed. Furthermore, the experiments showed that D binds less selectively than its analog A. Stable pairings were observed with A and G as complementary bases. This observation can be explained by possible binding patterns.

Unfortunately, I have not been able to synthesize the G analog homo-*C* nucleotide E and therefore I was unable to investigate an all-pyrimidine system.

Zusammenfassung

Die Entstehung des Lebens faszinierte die Menschheit und auch die Forschung seit jeher. Basierend auf der allgemein anerkannten *darwinschen Evolutionstheorie* versuchen unzählige Theorien die evolutionäre Entwicklung zu erklären. Speziell der Ursprung des Lebens, die Bildung und Eigenschaften der ersten "lebenden" Einheit erweckt unser Interesse.

In dieser Arbeit wurde versucht ein möglicher Vorläufers der heutigen Erbsubstanz DNA zu synthetisieren und auf seine chemischen und physikalischen Eigenschaften zu untersuchen. DNA besteht aus vier verschiedenen Nucleotidbausteinen, die wiederum in zwei Kategorien unterteilt werden: Purine und Pyrimidine. Eine Hypothese, aufgestellt von *Siegel* und *Tor*, schlägt ein System aus vier Pyrimidinbasen vor, aus dem die heutigen Nucleinsäuren entstanden sein könnten.[1] Die vier vorgeschlagenen Nucleobasen könnten alle aus ein und demselben Grundbaustein hervorgegangen sein, wobei die zwei kontemporären Pyrimidinebasen C und U, sowie zwei bis dato unbekannte Nucleotide D und E *Watson/Crick*–Basenpaarung eingehen würden. D und E, zwei Nucleotide mit einer exocyclischen amino-glykosidischen Bindung als Vorläufer von A respektive G müssten ähnliche Eigenschaften bezüglich der Stärke der Basenpaarung sowie Selektivität der Bindungspartner aufweisen (see *Figure 2*).

Um diese Voraussetzungen zu untersuchen wurde in dieser Arbeit eine Modellverbindung des Nucleotids D synthetisch hergestellt und in Oligonucleotide eingebaut. Auf Grund synthetischer Probleme wurde in der verwendeten Modellverbindung die amino-glykosidische Bindung durch eine stabilere Me-

thylenbrücke ersetzt. Das erhaltene homo-*C* Nucleotid sollte jedoch aus sterischer und elektronischer Sicht eine gute Imitation darstellen.

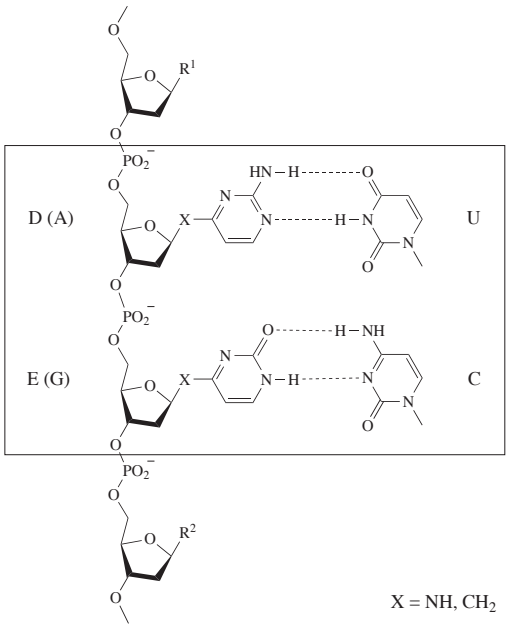


Figure 2: Vorschlag eines möglichen DNA Vorläufers.

Durch DNA-Schmelzexperimente von verschiedenen komplementären, doppelsträngigen Oligonucleotiden konnten die Basenpaarungseigenschaften von D qualitativ bestimmt werden. Es wurde gezeigt, dass D durchaus ähnliche Merkmale wie die Purinbase A besitzt und U in *Watson/Crick*–Weise binden kann. Allerdings zeigte D eine starke Abhängigkeit seiner *Nearest Neighborhood*, also der direkten Nachbarn in 5'- und 3'-Richtung. Vermutlich sind stark Dipol induzierende Basen in der direkten Nachbarschaft nötig um die flexible Nucleobase D in einer *stacked* Konformation zu stabilisieren und dadurch eine stabile Basenpaarung zu ermöglichen. Weiter zeigten die Experimente, dass homo-*C* Nucleotid D eine geringere Selektivität als sein Analog A aufweist und auch Bindungen zu A und G eingehen kann, was mit möglichen Bindungsstrukturen räumlich erklärt werden kann.

Leider gelang es uns nicht das G Analog homo-*C* Nucleotid E zu synthetisieren und somit ein reines Pyrimidinesystem zu untersuchen.

Theoretical Part

Chapter 1

Introduction

1.1 Chemical Evolution

The oldest procaryotic microfossils found are estimated to be 3–3.5 billion years old.[2] But what came before? How did such a diverse and complex structure evolve on a unwrought planet? The detailed process of prebiotic evolution is uncertain and must remain hypothetical because an unchallengeable proof for the primeval "living unit" is not possible. Nonetheless, supposition about the nature of this historical event motivates modern biochemical experiments, which lead to new discoveries.

First, one should define the term "Life," but opinions already differ on this highly philosophical question. From a biomolecular point of view, the most rudimentary definition of a "living unit" requires only the capability of reproduction, sexually or asexually. This definition requires a system that stores and imparts genetic information, and therefore maintains heredity. A more general definition includes the requirements of adaptation, metabolism, and growth and may involve, on a higher level, organization of several units. Adaptation means the ability to change over a period of time in response to environment, a quality which is fundamental to the process of evolution. Metabolism is necessary for growth, but, hypothetically, not implicitly re-

quired at an early stage of "life" if building blocks for replication are available.

All present-day living organisms are based on the same system, in which Deoxyribonucleic acids (DNA) contain the genetic information and Ribonucleic acids (RNA) are transcribed from DNA and carry the code of single genes. This information is then translated into peptides, the molecular machines of cells. These three biopolymers are presumably intermediates of the development toward present-day cellular mechanisms and one of these was possibly even the first biopolymer. It is crucial to find out which biopolymers or building blocks could have been available and stable under prebiotic conditions and which evolutionary pathways could have led to the other elements of the current system. The search for a prebiotic hereditary material can, therefore, be limited to two questions:

– which of the following biopolymers was first: DNA, RNA or peptides?
– what are possible precursors of that biopolymer?

This discussion is hypothetical and is based on model systems because questions regarding conditions on early earth and the point in time when life evolved must remain open. The latter can be only roughly estimated to be ∼4 billion years ago and this makes it difficult to determine the environment at this time. Since the pioneering work of *Haldane* and *Oparin*, the prebiotic soup theory has dominated the theories about how life emerged on earth.[3,4] According to a modern version of this theory, organic compounds accumulated in the primordinal oceans, underwent polymerization and generated increasingly complex macromolecules.[5] In general, the scientific community agrees on assumed conditions like moderate pH, reducing atmosphere (lack of O_2) and aqueous ambience but one crucial controversial subject is the temperature. One theory suggests that life originated at hydrothermal vents, which is based on the fact that thermophilic microorganisms, which grow at temperatures between 80 °C and 110 °C, are claimed to be the oldest organisms known so far.[6] But many experts doubt this theory due to temperature

lability of biomacromolecules.[7]

Another motivation for intensive discussions is the dilution problem. Polymerization of available building blocks implies a fairly high concentration of those monomers, a prerequisite which was not fulfilled in primordinal oceans. Accumulation in small puddles or enrichment on mineral surfaces is considered to solve this problem.[8] Minerals can also serve as templates for polymerization and crystallization.[9] A second solution for the dilution problem may be compartments, enclosure of simple building blocks or even simple metabolisms in micelles.[10]

The importance of minerals is also a contended subject. In addition to their ability to absorb organic compounds, they can also stabilize them in diverse fashions (e.g. ribose is stabilized by borates[11]) or they can catalyze countless reactions. *Wächtershäuser* proposed a highly regarded theory about surface metabolism.[12,13] He suggested that the first primitive metabolic system was a series of reactions based on monomeric organic compounds made directly from simple constituents (CO, CO_2, NH_3, H_2S) in the presents of metal sulfide catalysts and hot magmatic exhalations.[14]

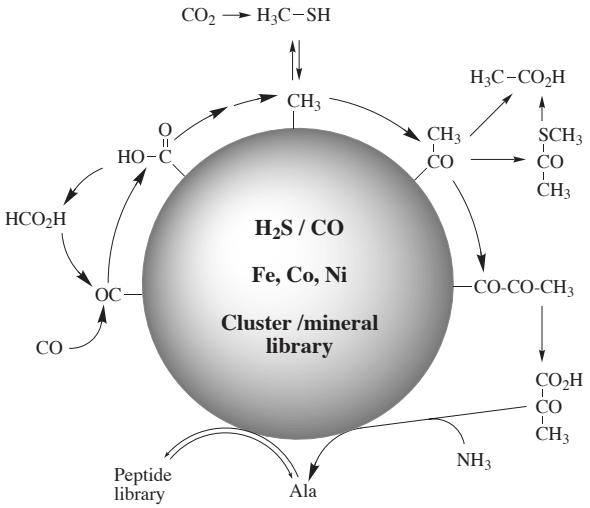


Figure 1.1: Metabolism, catalyzed by sulfide clusters.

Even though *Wächtershäuser's* proposal describes only a metabolic circle without any considerations about hereditary material, it could be an important piece of the puzzle, helping us to explain how building blocks or even biopolymers were synthesized under prebiotic conditions.

Another approach toward the formation of building blocks and biopolymers is the famous experiment by *Urey* and *Miller* in 1953.[15,16] They tried to simulate the atmosphere on primitive earth and exposed a mixture of H_2O, CH_4, NH_3, H_2 and CO to electric discharges. Out of the product mixture, they were able to isolate a series of organic compounds like formic acid, acetic acid, glycine, alanine, urea *et cetera*. This discovery was a milestone in prebiotic science, but some doubts were raised with respect to its accuracy. Probably, the simulated conditions were too mild compared to locally distributed rough conditions in the primal atmosphere. Additionally, the UV sensitivity of NH_3 and CH_4 could have also constituted a problem.

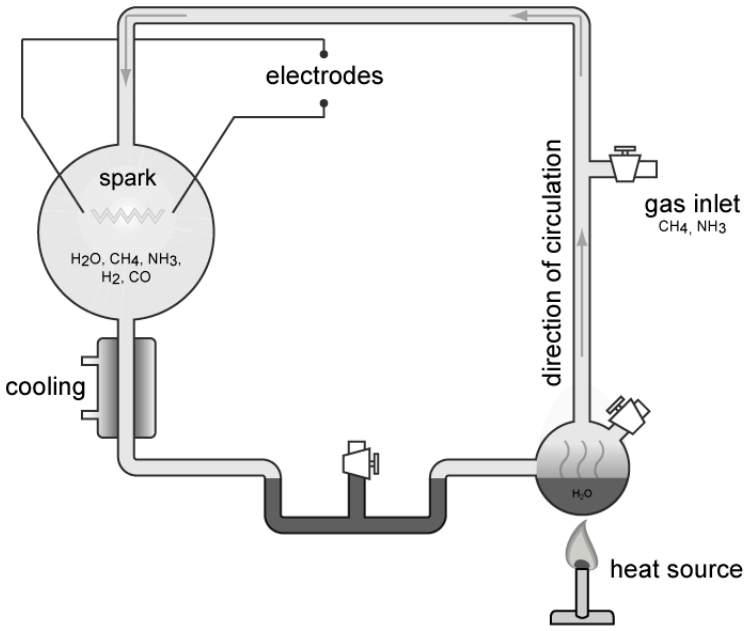


Figure 1.2: Apparatus of the *Miller-Urey* experiment.

Another crucial problem of chemical evolution science arose in *Miller's* discovery because he detected all amino acids as racemic mixtures. The origin

of nature's *homochirality*, why proteins and nucleic acids are composed of predominantly L-amino acids and D-sugars is still uncertain. Several approaches try to explain homochirality in consequence of asymmetric autocatalysis, asymmetric catalysis by prebiotic amino acids which were enriched in one of the enantiomeric forms, or enantiomeric adsorption on chiral mineral surfaces like quartz.[17–21] Another theory by *Mills* proposes an early racemic system, which evolved to a chiral one due to kinetic advantages of the latter.[22] It is known for example, that oligomerisation of monomers of the same handedness polymerize faster. Summarized, all these theories point to a spontaneous resolution of one of the enantiomers as likely.

1.1.1 The RNA World hypothesis

As mentioned above, searching for the prebiotic hereditary material leads to the conclusion that either DNA, RNA, or peptides are possible candidates. Its special function, and its importance in the cell's metabolism, turned scientist's attention to RNA. Prestigious scientists like *Francis Crick*, *Manfred Eigen*, *Walter Gilbert* and *Leslie Orgel* are regarded as the founders of the *RNA World hypothesis*. Several facts count for RNA being the starting point of the chemical evolution of our cellular metabolism. First, RNA is the connection between genetic material and metabolism and, second, RNA was found to have the ability to contain genetic information and some enzyme-like catalytic activities. It's obvious that RNA can hold genetic information in the same fashion as DNA but the discovery of several enzymatic activities in the 80's of the last century was very surprising. *Altman* et al. discovered RNA that acts as a coenzyme in ribonuclease[23–25] and *Cech* et al. found ribosomal RNA containing a self-splicing exon.[26–28] With these new insights, one can hypothesize that RNA is able to participate in self-replication and recombination.[29] Accordingly, an RNA World can be contemplated that contains only RNA molecules that served to catalyze the synthesis of themselves,[30] and that developed a range of enzymic activities by using cofactors.[31] Later,

RNA could have begun to synthesize proteins followed by the appearance initially of single stranded DNA. Later, evolution of the complementary double strand allowed stable linear information storage as well as error correction. It seems reasonable to conclude that the RNA World did exist at some time but many doubts have been raised about whether or not the primordial hereditary material could have been RNA. In order to be a potential prebiotic biopolymer, it should be stable under primordial conditions, but, as explained in *section 3.3.2.7*, RNA is labile to hydrolysis and to UV background radiation.[32] Additionally, ribonucleic acid synthesis under prebiotic conditions has not been successful in the laboratory so far.[33]

1.1.2 Possible Precursors of DNA/RNA

Due to the stability problems of RNA which were mentioned in *section 1.1.1* and the consequential fact that the precise chemical nature of the original genetic material presumably differed from present-day RNA the search continues for a possible precursor to RNA or DNA. Such a progenitor requires to be stable and synthetically accessible under prebiotic conditions. A series of candidates were suggested in the literature but only the most reasonable examples should be mentioned here. Concerning evolutionary mutations within the nucleic acid structure, three parts can be examined separately: backbone, sugar and base. As a backbone modification, *Miller* and *Orgel* suggested *Peptide Nucleic Acid* (PNA) as a possible precursor of RNA.[34] PNA consists of AEG (*N*-(2-amino- ethyl)glycine) units, which were shown to be available from CH_4, N_2, NH_3 and H_2O in electric discharge reactions.[35] PNA binds DNA and forms double and triple helical structures that are related to the *Watson-Crick* helix.[36,37]

Another candidate was first prepared by *Ueda* in 1971: *Glycerol nucleic acid* (GNA).[38] This analogue also showed pairing to DNA and RNA as well as self-pairing. Although it is composed of an acyclic three carbon backbone unit, GNA exhibits stronger *Watson-Crick* pairing in duplex than its natu-

ral counterparts. Furthermore, the building blocks should have been easily available under prebiotic conditions.[39–41]

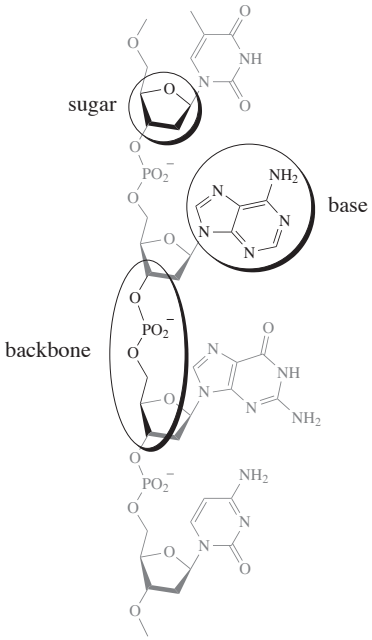
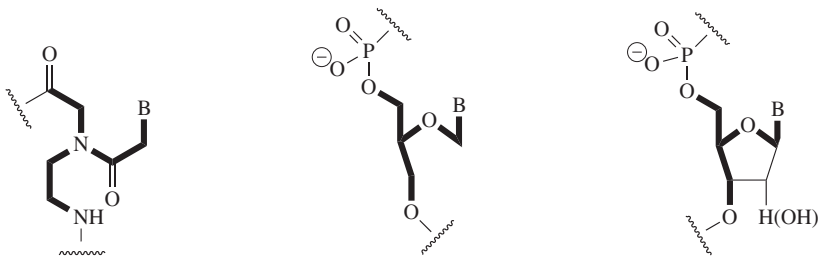


Figure 1.3: Points of attack for possible mutations.



(a) PNA *(b)* GNA *(c)* DNA(RNA)

Figure 1.4: Possible backbone precursors in comparison with DNA(RNA).

As seen in *Figure 1.4*, PNA units are, in contrast to present-day nucleic acids, connected by amides. Another crucial question of chemical evolution sciences is, why nucleosides are connected by phosphodiester groups? In 1987, *Westheimer* discussed the advantages of phosphodiesters holding sugar units together in RNA and DNA[42] and a pioneering study by *Usher* dealt the question of why the junction is positioned between carbons 3' and 5' instead

of 2' and 5'. He found experimentally the (2'→5') junction to be more susceptible to hydrolytic strand cleavage.[43,44]

Albert Eschenmoser systematically studied many possible sugar modifications.[45–49] He asked why nature developed the structure type of ribofuranosyl nucleic acids, rather than some other sugar family, and prepared nucleic acid systems with several different sugar moieties. He found the capability of *Watson-Crick* base-pairing to be a widespread property among potentially natural nucleic acid alternatives. The most promising systems were the so-called homo-DNA and TNA.[45,49,50]

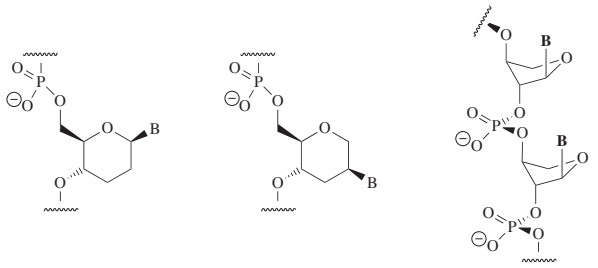


Figure 1.5: Structure of two different Homo-DNA's and TNA.

He uses the term "homo-DNA" for an artificial oligonucleotide system differing from DNA in an expanded six-membered pyranose ring by an additional methylene group. Due to its chemical composition, homo-DNA was never supposed to be a potentially natural nucleic acid alternative but it showed much stronger *Watson-Crick* base pairing than DNA itself, presumably a consequence of the higher rigidity of pyranose rings, compared to those of furanose.[51] An constitutional isomer of homo-DNA was investigated by *Van Aerschot* et al. (see *Figure 1.5*, middle). He linked the base to the sugar at carbon 2' and prepared oligonucleotides, which show strong *Watson-Crick* base pairing and additional cross pairing to RNA.[52]

Even more surprisingly was the discovery of pairing abilities of TNA ((L)-α-THREO-furanosyl-(3'→2')-oligonucleotides). *Eschenmoser* et al. prepared these tetrose sugar units, containing only five covalent backbone bonds per

repeating unit, which show efficient base pairing similar to canonical RNA with regard to specificity, strand orientation and pairing strength.[50] In addition, (L)-TNA oligonucleotides are capable of cross-pairing with RNA and DNA, which is remarkable when one considers the constitutional and conformational differences between these polymers. Furthermore, TNA is much more stable toward hydrolytic cleavage of the phosphodiester linkage than RNA and could be presumably formed under prebiotic conditions from glycolaldehyde.[53]

A third site for evolutionary modifications is the base. The instability of nucleobase C (see *section 3.3.2.7*), which shows on the geologic timescale a very short half-life for the decomposition, even at 0°C, raises the question of whether it would have been a suitable base for the first genetic material. Suggestions for a prebiotic harbinger of a present-day complementary base system are rare, but a series of alternative bases and base pairs have been proposed. For example, *Benner* et al. suggests isoguanine (isoG) and isocytosine (isoC), diaminopurine (amino A) and U or pseudo-diaminopyrimidine (pseudo D) and xanthine.[54] Another proposal was published by *Miller* et al.: A and 1,2,4-triazole-3,5-dione (urazole).[55]

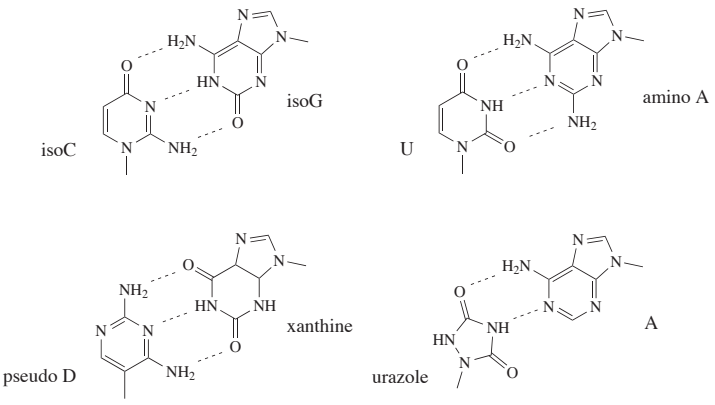


Figure 1.6: Alternative base pairs for GC.

Wächtershäuser proposed an all-purine system (*Figure 1.7*).[56] His assumption is motivated by the fact that purines can be oligomerized on a com-

plementary template in the absence of polymerase. But pyrimidines cannot, probably because of low stacking energies.[57] Furthermore N^3 bound purines have been found in prokaryote and eukaryote cells.[58,59]

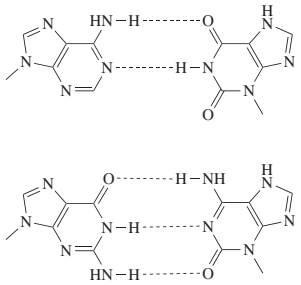


Figure 1.7: Proposed binding pattern of an all-purine system.

A completely opposite proposal was posted by *Siegel* and *Tor*: an all-pyrimidine precursor (*Figure 1.8*).[1] They postulate a possible precursor to be compatible with the present system, but less fit for the environmental pressures that motivates mutational evolution. It should be isosteric with the modern code and accessible from a single progenitor heterocycle. Such a progenitor could be 2,4-diaminopyrimidine, which was presumably available on prebiotic earth and can be converted by hydrolysis to C and U (see *section 3.4.1*). By unequal formation of the glycosidic bond, a four letter code containing only these three pyrimidine bases can be assumed as outlined in *Figure 1.8*. A detailed discussion about this system will follow in the next chapter.

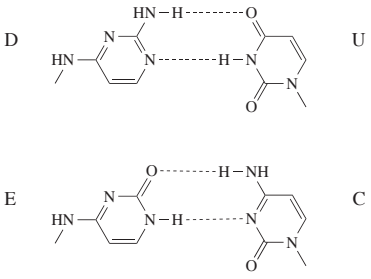


Figure 1.8: Proposed binding pattern of an all-pyrimidine system.

Chapter 2

The Project

As mentioned before, *Siegel* and *Tor* published a proposal of an all-pyrimidine precursor of DNA or RNA (see *section 1.1.2*, *Figure 2.1*).[1]

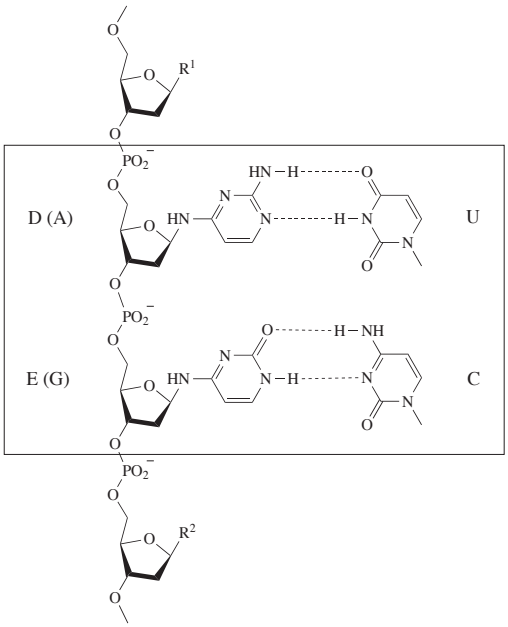


Figure 2.1: DNA/RNA precursor based on pyrimidines.

Based on the knowledge that cytosine (C) is the least stable of the four present day nucleobases and undergoes spontaneous hydrolysis to uracil (U) (see *section 3.3.2.7*) one can assume cytosine to be a hydrolysis product of 2,4-diaminopyrimidine (D) (*Figure 2.2*). With these three pyrimidines, one

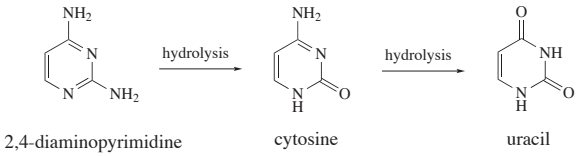


Figure 2.2: Hydrolysis cascade of 2,4-diaminopyrimidine.

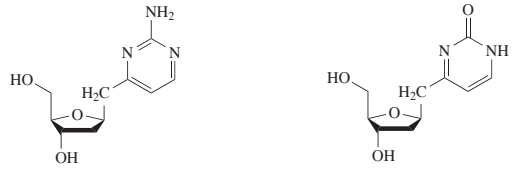
can draw a pattern consisting of the known nucleosides C and U and two unknown nucleosides building the glycosidic bond through an exocyclic amino group. Nucleoside D would be isosteric to adenosine (A) and E would be isosteric to guanosine (G), each forming a motif with two hydrogen bonds. From an evolutionary point of view, these exocyclic amino nucleosides (EANs) could be harbingers of A and G, respectively. EANs are not only hypothetical constructs. They are also known in nature as degradation products of purines under oxidative pressure[60] or natural metabolites.[61] Especially the fact that purines degrade under oxidative stress is of interest. Assuming a strongly oxidative atmosphere, purines would not be existent in the absence of a repairing system, but EANs would. Therefore its quite reasonable to propose EANs as harbingers of present-day purines. A few examples of naturally occurring EANs will be illustrated in *section 3.4*.

The aim of this thesis is to investigate and possibly support this hypothesis chemically. Therefore, building blocks should be synthesized to investigate their stability and physicochemical properties. Additionally, oligonucleotides containing this artificial nucleotides are of interest to determine their ability of base pairing in a *Watson-Crick* mode.
A summary of conditions that an all-pyrimidine system should fulfill to be a possible precursor of the contemporary hereditary material is listed bellow.[51] Some of these conditions are difficult to prove and will not be a target of this project; however, they should be investigated in the future.

– accessible from simple molecules available on primitive earth
– available by the same type of potentially natural chemistry that allows

the structure of RNA to be derived from ribose
– capable for informational base pairing in the *Watson-Crick* mode
– able to self-replicate non-enzymatical under potential conditions
– able to express a chemical phenotype in the given environment

Preliminary studies performed in our laboratory[62,63] and by others[64] have shown EANs to be susceptible to anomerisation, deglycosylation, and isomerisation. The presence of strong electron-withdrawing substituents on the pyrimidine ring can diminish these problems, but such a modification would obviously change pyrimidine's electron density and therefore its ability for base pairing. To avoid this restriction, we decided to switch to carbon analogues, where the glycosidic bond is formed through a bridging methylene group instead of an exocyclic amino function. *Greenberg* et al. showed such *C*-nucleosides to be useful models.[65] These so-called homo-*C* nucleosides should be thermodynamically more stable than the corresponding EANs, sterically comparable and in addition not dramatically different concerning electron density, as shown by *Berstis*.[66]



homo-*C* nucleoside D homo-*C* nucleoside E

Figure 2.3: Homo-*C* analogs of nucleosides D and E.

Chapter 3

State of Knowledge

3.1 Historical Overview

At the time when *Charles R. Darwin*[67] (1859) and *Gregor Mendel*[68] (1866) published their famous theories, the chemical knowledge about processes in living organisms was very rudimentary and limited to the physical dimension of cells. General opinion, of the time, considered genes to be proteins, because proteins were the only biochemical building blocks one estimated to exhibit the necessary specificity. *Ernst Haeckel*[69] proposed in 1866 that the *Nucleus* contains the factors responsible for the transmission of hereditary traits and only three years later, in 1869, *Friedrich Miescher* discovered DNA by observing a precipitate from a leucocyte extract under acidic conditions.[70,71] He noticed that this leucocyte extract, obtained from pus on fresh surgical bandages, contained a substance which precipitated when acid was added and dissolved again when alkali was added. Due to its presence in the nuclei, *Miescher* termed the enigmatic compound "nuclein". Later he was able to show that nuclein, unlike proteins, lacked sulfur but contained a large amount of phosphorus.[72] Although neither the structure nor the exact function of this new substance was yet known, *Walther Flemming*[73] described the morphology and behavior of the chromosomes during mitosis and, only one decade later, *Theodor Boveri*[74] made the case that the chromosomes not

only harbor the genetic information, but that individual chromosomes carry different parts of the hereditary material.

Within two decades, from 1909 to 1929, *Phoebus Levene* characterized RNA and DNA, identified the four DNA builing blocks A, C, G and T and formulated the *tetranucleotide hypothesis*, based on the fact that he had found the four building blocks in equal ratios. On the basis of this hypothesis he postulated a monotonously repeating sequence, which consequently would not be able to have genetic function (see *Figure 3.1*).[75]

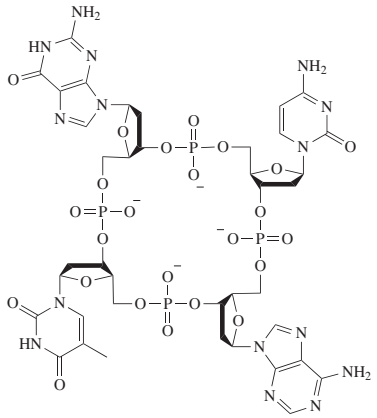


Figure 3.1: Hypothetical tetranucleotide, as proposed by *Levene*.

Approximately at the same time, *Frederick Griffith* postulated the *transforming principle*.[76] He observed that a mixture of heat-killed virulent S pneumococci (*Diplococcus pneumoniae*) and non-pathogenic R pneumococci is still contagious, both *in vivo* and *in vitro*. He concluded that a *transformation* of R to S occurred, induced by the disarmed S pneumococci. The nature of this transformation remained unclear until *Oswald Avery*, *Colin MacLeod* and *Maclyn McCarty* showed DNA to be the *transforming principle*, in 1944.[77] They reported this principle to have all physical and chemical properties of DNA. It contained no detectable proteins, was unaffected by trypsin, chymotrypsin and ribonuclease, but was completely inactivated by DNase.

Unfortunately, *McCarty's* insight remained unappreciated until 1949 when *Erwin Chargaff* determined the accurate ratios of nucleobases in DNA and disproved the tetranucleotide hypothesis. Thereby he indicated that DNA could be complex enough for a potential genetic function;[78] however, decisive evidence was found by *Ralph Brinster* et al. three decades later![79] *Brinster's* team proved that DNA is also the transforming principle in eukaryotes by injection of DNA bearing the gene for rat growth hormone into *in-vitro* fertilized mouse eggs and implantation of those into the reproductive tracts of foster mothers. The resulting *transgenic* mice had high levels of rat growth hormone in their serum and reached nearly twice the size of their natural littermates.

From the perspective of structural analysis of DNA, the first momentous breakthrough was accomplished by the two teams *Franklin–Wilkins* and *Watson–Crick*, with the clarification of the helical structure.[80–82] *Franklin* and *Wilkins* demonstrated that DNA has a regularly repeating helical structure, and *Watson* and *Crick* discovered the molecular constitution of DNA. They found the later to be a double helix in which A always pairs with T, and C always with G. Additionally, *Crick* proposed four years later the translation of information in DNA into proteins through RNA and he speculated that one amino acid in a protein is always coded by three bases in DNA.[83]

Two other milestones in the history of DNA should be mentioned here:
1) In the late 1970s several methods to sequence DNA were developed by *Frederick Sanger*, *Allan Maxam* and *Walter Gilbert*.[84,85]
2) In 2001 the complete sequence of the human genome was puplished.[86]

3.2 Biological Function Of Nucleosides, Nucleotides and Nucleic Acids

Nucleotides are biologically ubiquitous substances. In addition to their central role in both storage and expression of genetic information, they participate in many biochemical processes.[87] In this section only a short overview of biological functions of nucleosides should be outlined. In respect of the widespread biological activities of nucleosides, analogs, as synthesized in this thesis, are always of interest to be tested in biochemical studies. *Section 4.7* will summarize possible application fields of nucleoside analogs and therefore the biochemical background should be discussed here.

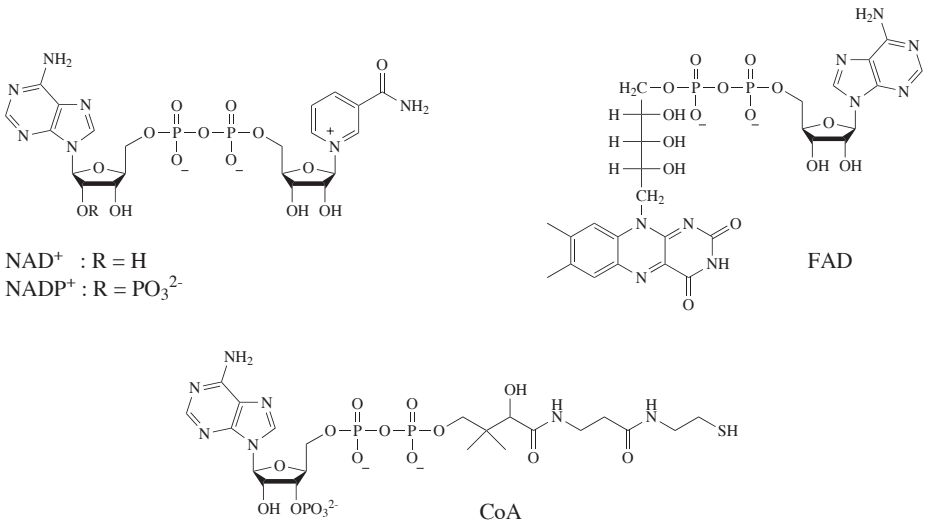


Figure 3.2: Nucleotide coenzymes NAD^+, $NADP^+$, FAD and CoA.

The most widespread and probably most investigated derivatives of nucleosides are ATP or nucleoside triphosphates in general. Since ATP was discovered in 1929 by *Karl Lohmann*[88] it has been recognized as an energy source for many important reactions and cells have complex mechanisms for its synthesis. More recently, ATP (and/or ADP) was found to influence many biological processes like platelet aggregation, vascular tone, neurotransmission, cardiac function and muscle contraction.[89] Nucleotide derivatives, such

as nicotinamide adenine dinucleotide (NAD^+, $NADP^+$), flavin adenine dinucleotide (FAD) or coenzyme A (CoA), are known to serve as coenzymes (see *Figure 3.2*).

In 1982, *Cech* et al. demonstrated that some RNAs undergo autocatalytic rearrangements, also called *splicing*.[26] They drew the conclusion that RNA cleavage and ligation activity are intrinsic to the structure of the molecule. Furthermore, RNA was also found to act as a coenzyme in RNA-dependent DNA polymerase, for instance.[90,91]

3.3 DNA Structure

3.3.1 Nucleosides and Nucleotides

Nucleic acids can be divided into two main classes: Ribonucleic acids (RNA) and Deoxyribonucleic acids (DNA). Their building blocks, called nucleotides, are composed of a nucleobase, a sugar moiety and a phosphate group. They differ foremost in the sugar moieties, which are the pentose sugars D-ribose in RNA and 2-deoxy-D-ribose in DNA.

Both classes of nucleic acids consist of four building blocks, which can again be subdivided into two categories: purines and pyrimidines. The purine nucleotides adenosine (A) and guanosine (G) as well as the pyrimidine nucleotide cytidine (C) appear in both RNA and DNA. By contrast, the remaining pyrimidine nucleotide is found in RNA as uridine (U) and in DNA as thymidine (T, 5-methyluridine).

The pentose sugar is locked in a five-membered furanose form, linked to the nucleobase at C1'. This so-called *glycosidic* bond is on the same side of the sugar as the C4'–C5' bond and is termed to be in a β-configuration (R vs. α-configuration, which is S). The phosphate group may be bonded to the 5'-(5'-nucleotide) or to the 3'-hydroxy group (3'-nucleotide). In the absence of a phosphate group the unit is called a nucleoside. In *Figure 3.3*, all of these

five nucleosides are shown, including atom numbering and nomenclature of nucleobases (e.g. adenine) and nucleosides (e.g. adenosine), following the recommendation of *IUPAC* CBN.[92]

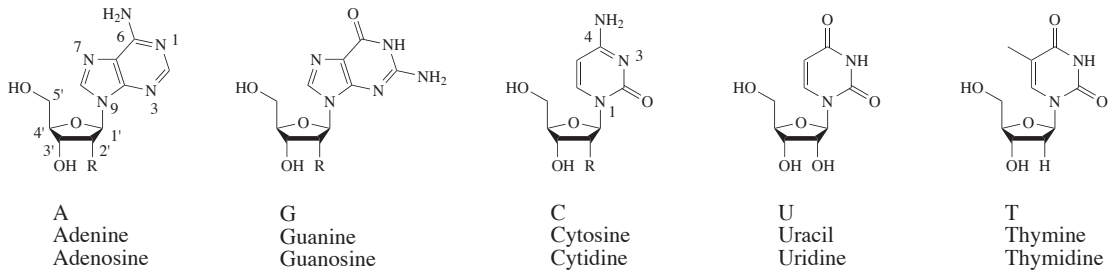


Figure 3.3: Structures of the five major nucleosides in DNA and RNA.

3.3.2 Physical Properties of Nucleosides and Nucleotides

3.3.2.1 Protonation

For a better understanding of the detailed structure and hydrogen bonding ability of DNA, RNA, and their constituents, it is crucial to know exactly the sites and thermodynamic quantities associated with the interaction with protons. For example, protonation plays an important role in determining the base form available for base pairing and recognition in macromolecular structures because a protonated nucleobase has obviously other hydrogen bond acceptor and donor abilities than it's neutral counterpart.[93]

As seen from the pK_a values all bases are uncharged in a pH range of 5–9 (*Table 3.1*). The phosphodiester group has a pK_a value of $\sim$ 2.6 (H_3PO_4: $pK_{a_1} = 2.15$). In former times, it was stated that ionization of A, G and C occurs at the exocyclic amino group, which is comprehensible due to the greater electron density on exocyclic, rather than ring nitrogens. However, it was later concluded that the most basic site is not necessarily determined by the highest electron density, but rather by the conditions in the protonated state.[95] That means in other words, the better the protonated species is

Table 3.1: pK_a values for bases in deoxy-nucleosides and -nucleotides (conc. 5^{-5} – 10^{-5} M, 20 °C, zero salt conc.).[94]

Base, site of protonation		Nucleoside	5'-Nucleotide	3'-Nucleotide
Adenine	N^1	3.5	3.9	3.7
Guanine	N^7	3.3	–	–
	N^1	9.4	10.0	9.8
Cytosine	N^3	4.2	4.6	4.4
Thymine	N^3	9.9	10.5	–
Uracil	N^3	9.4	10.1	10.0

stabilized by resonance structures, the more basic is the corresponding protonation side. This finding led to the conclusion that in adenine, for example, N^1 is the most likely site of protonation, which was confirmed by calculations, X-ray and NMR.[96–98]

For the sake of completeness it should be mentioned here that the 2'-OH of ribonucleosides and their corresponding 3'-monophosphates were also investigated. The pK_a values of the 2'-OH group of A, G, U and C were found to be 12.2, 12.5, 12.6 and 12.5, respectively. In addition, the pK_a values for the 3'-monophosphates of AMP, GMP, UMP and CMP are 13.4, 13.5, 13.6 and 13.6, respectively.[99]

3.3.2.2 Tautomerism of Bases

Nucleobases, regardless if present as nucleotides, nucleosides or as simple heterocycles without a sugar entity, exist in solution in an equilibrium of tautomers. The concentration of the preferred tautomer usually exceeds that of the others by a factor of 10^3–10^6.[100] This predominance of one major tautomer is absolutely essential, because tautomers show different hydrogen bonding donor/acceptor properties (D/A) and can, therefore, form

non-Watson/Crick base pairs (*Scheme 3.4*). Such non-Watson/Crick base pairs could lead to point mutations if they are not detected by repair enzymes.[101–103] The predominance of one major tautomer is also necessary to ensure a correct recognition of complementary bases, which is the quintessence for DNA replication and transcription.

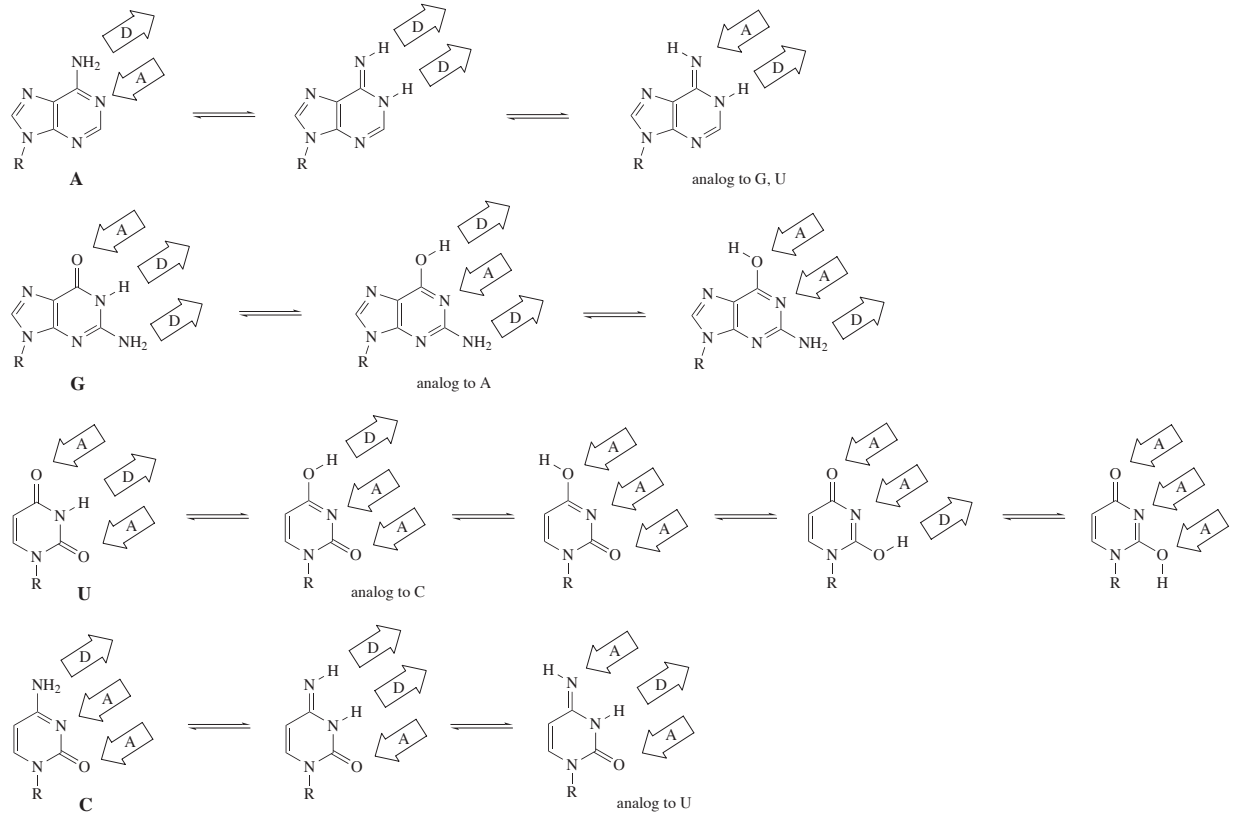
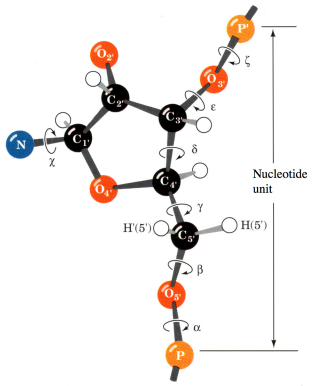


Figure 3.4: Keto-enol and amine-imine tautomerism of A, G, U and C.

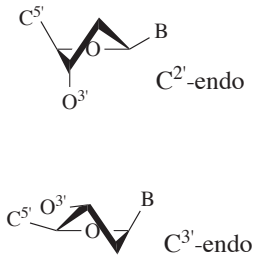
3.3.2.3 Sugar-Phosphate Chain Conformation

Considered superficially, one could expect nucleic acids to be very flexible, due to all possible bond rotations defined in *Figure 3.5a* as torsion angles α, β, γ, δ, ε, ζ and χ. But, these seven torsion angles are subjected to many intramolecular restrictions, which make the structure much more rigid than estimated. Starting the analysis with χ, the torsion angle of the glycosidic bond, it is obvious that only two orientations of the base are sterically rea-

sonable: *syn* and *anti* conformations (see *Figure 3.6*). In most double helixes, only *anti* conformers are found in both purine and pyrimidine nucleotides. An exception is *Z*-DNA, where purines adopt the *syn* orientation.



(a) Torsion angles of a nucleotide unit[87]



(b) Preferred *endo* conformations

Figure 3.5: Degrees of freedom in sugar and phosphate backbone.

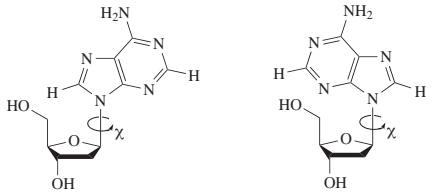


Figure 3.6: *anti* and *syn* conformer of dA.

The flexibility of the sugar-phosphate backbone is limited as well. As shown by *Sundaralingam*,[104] bond rotations in the backbone chain are confined by covalent stiffness of the ribose as well as non-covalent interactions of the phosphate group. Nevertheless, the sugar-phosphate angles in double helix are conformationally quite strain-free.

Only two envelope conformations are found in nucleosides and nucleotides: $C^{2'}$-*endo* and $C^{3'}$-*endo* (*Figure 3.5b*). *Endo* refers to the case where the $C^{2'}$ or $C^{3'}$ sticks out of the plane to the same side as $C^{5'}$. According to quantum-chemical results, these two conformations undergo rapid equilibrium with a

barrier of ca. 24 kJ/mol.[105] *B*-DNA occurs only as $C^{2'}$-*endo* ribose and *A*-DNA only as $C^{3'}$-*endo*. In contrast *Z*-DNA pyrimidine nucleotides are present as $C^{2'}$-*endo* riboses and purine nucleotides as $C^{3'}$-*endo*.

3.3.2.4 Primary and Secondary Structure of DNA

Already twenty years before the helical structure of DNA was found, *Klein* and *Thannhauser* described the primary structure of nucleic acids.[106,107] They established that the nucleosides are joined in a 5'→3' manner by a phosphodiester junction and that there are neither 5'–5' nor 3'–3' linkages. This observation is evidence that the uniqueness of a DNA sequence resides only in the succession of its bases.

Watson and *Crick*[81] deduced the helical secondary structure of *B*-DNA in 1953 from fiber diffraction patterns and in the same year *Franklin* and *Gosling* reported the discovery of *A*-DNA.[108] *Franklin* and *Gosling* also deduced from their observation that a conversion between *B*-DNA and *A*-DNA should be a reversible conformational change, which is dependent on the relative humidity and the nature of the present cations. In 1979, a third example of secondary structure was discovered. *Wang* and *Rich* found a left-handed double helix and termed it *Z*-DNA.[109] *Table 3.2* summarizes the structural characteristics of these three main occurrences of DNA.

At the time when *Watson* and *Crick* determined the helical structure of DNA they had only cell extracts and, therefore, crystals of poor quality available. It took another three decades until the first X-ray crystal structure at near-atomic resolution was published. In 1980, *Dickerson* et al. were able to measure an X-ray with 1.9 Å resolution of an artificially synthesized oligonucleotide (dGlc[CGCGAATTCGCG]).[110] This was the first single-crystal structure analysis of more than a complete turn.

Table 3.2: Structural characteristics of *A*-, *B*- and *Z* -DNA.[87]

	***A*-DNA**	***B*-DNA**	***Z*-DNA**
Helical sense of rotation	right-handed	right-handed	left-handed
Diameter	∼ 26 Å	∼20 Å	∼18 Å
Base pairs per helical turn	11.6	10	12 (6 dimers)
Helical twist per base pair	33°	36°	60° (per dimer)
Helix pitch (rise per turn)	28 Å	34 Å	45 Å
Helix rise per base pair	2.6 Å	3.4 Å	3.7 Å
Base tilt normal to the helix axis	20°	6°	7°
Major groove	narrow and deep	wide and deep	flat
Minor groove	wide and shallow	narrow and deep	narrow and deep
Sugar pucker	$C^{3'}$-*endo*	$C^{2'}$-*endo*	$C^{2'}$-*endo* for pyrimidines $C^{3'}$-*endo* for purines
Glycosidic bond	*anti*	*anti*	*anti* for pyrimidines *syn* for purines

3.3.2.5 DNA Duplex Stabilization

Since *Watson* and *Crick* presented their structural proposal of DNA,[81] the non-covalent interactions which influence the oligonucleotide duplex have been investigated intensively.[111,112] Some forces, including base stacking and dipole-dipole interactions are stabilizing and other forces destabilize the duplex, such as electrostatic repulsion of the phosphates. Overall, DNA double

helices are strongly favored enthalpically and are disfavored nearly as much entropically.[113] Rapidly, it became clear that hydrogen bonding is essential for specific base pairing and ensures the reliability of DNA replication, but makes only a small contribution to duplex stabilization.[105] The main factors for the latter are base stacking, dipole-dipole as well as hydrophobic interactions. In 1964, *Steward* and *Jensen* described base stacking of adenine by X-ray crystallography[114] and showed a partial overlapping, which is characteristic for the association of bases in crystals and in double helical nucleic acids. Several noncovalent forces play a role in base stacking. Van der Waals dispersive forces (dipole–induced dipole and induced dipole–induced dipole attractions), permanent electrostatic effects of interacting dipoles, and solvation effects influence the stacking stability.[115] In general, the purines stack more strongly than the pyrimidines, presumably because they have larger surface area and greater polarizability. *Kool* et al. measured the stacking free energy of natural bases as dangling ends in a short DNA duplex with a C–G base pair at the end to be 0.5, 0.6, 0.7 and 1.0 kcal/mol for C, T, G and A, respectively.[116] But it needs to be considered that this stacking energy is strongly dependent on its nearest neighboring bases, an influence called the *Nearest-Neighbor Effect*.[117] Recent studies have established that the stacking of DNA and RNA bases is a strong contributor to the overall stabilization of the double helix, and may be the dominant one.[112]

As already mentioned, hydrogen bonding does not play a crucial role in DNA stabilization, but dipole–induced dipole interactions do. The strong dipole of C induces another strong dipole in its easily polarizable complementary *Watson/Crick*–base G, for example, and leads to a binding between the two stands.[118] The same effect occurs in T–A base pairs in a diminished manner. Ionic interactions affect the duplex stability significantly. Electrostatic repulsion of the anionic phosphate groups destabilizes the double helix but can be degradated by cations, whereas divalent cations like Mg^{2+}, Mn^{2+} or Co^{2+} shield the phosphate groups much more effective than Na^{+} or K^{+}.[119,120]

3.3.2.6 Hydrogen Bonding Motifs

In addition to the *Watson-Crick* base pairs, 26 alternative mispairs for A, G, U (T) and C, involving at least two hydrogen bonds, are structurally possible.[105] In this chapter, we will only describe the biochemically most relevant *non-Watson-Crick* base pairing motifs.

In 1963, *Karst Hoogsteen* heated a solution of adenine and thymine and then let it cool slowly to form crystals. He found that in his crystals, adenine and thymine did not form hydrogen bonds as proposed by *Watson* and *Crick*.[121] Instead, they formed two hydrogen bonds, that involved the N^7 atom of the purine ring rather than the N^1 atom. The *Hoogsteen* geometry is the most favourable one for A–T base-pairs in solution, but not in double helices. G–C base pairs can only form *Hoogsteen* base pairs if the cytosine N^3 is protonated. *Hoogsteen* base pairing can be seen in three-dimensional structures of transfer RNA, in triple-stranded helices when a third stand winds around a duplex that is assembled in the *Watson-Crick* pattern or in G-quadruplexes, where this alignment allows assembly of planar quartets which are composed of stacked associations of *Hoogsteen* bonded guanines (see *Scheme 3.8*).[122] *Watson-Crick* and *Hoogsteen* pairing can also occur in a *reversed* manner, where one base is rotated by 180° with respect to the other. Reversed *Hoogsteen* is found to be especially important in tRNA function.[123]

Another important hydrogen bonding motif was found by *Francis Crick* in 1966: the *Wobble* configuration.[124] The G–U *wobble* base pair is a fundamental unit of RNA secondary structure that is present in nearly every class of RNA. It has thermodynamic stability that is comparable to *Watson-Crick* base pairs and is nearly isomorphic to them. Therefore, it often substitutes for G–C or A–U base pairs.[125]

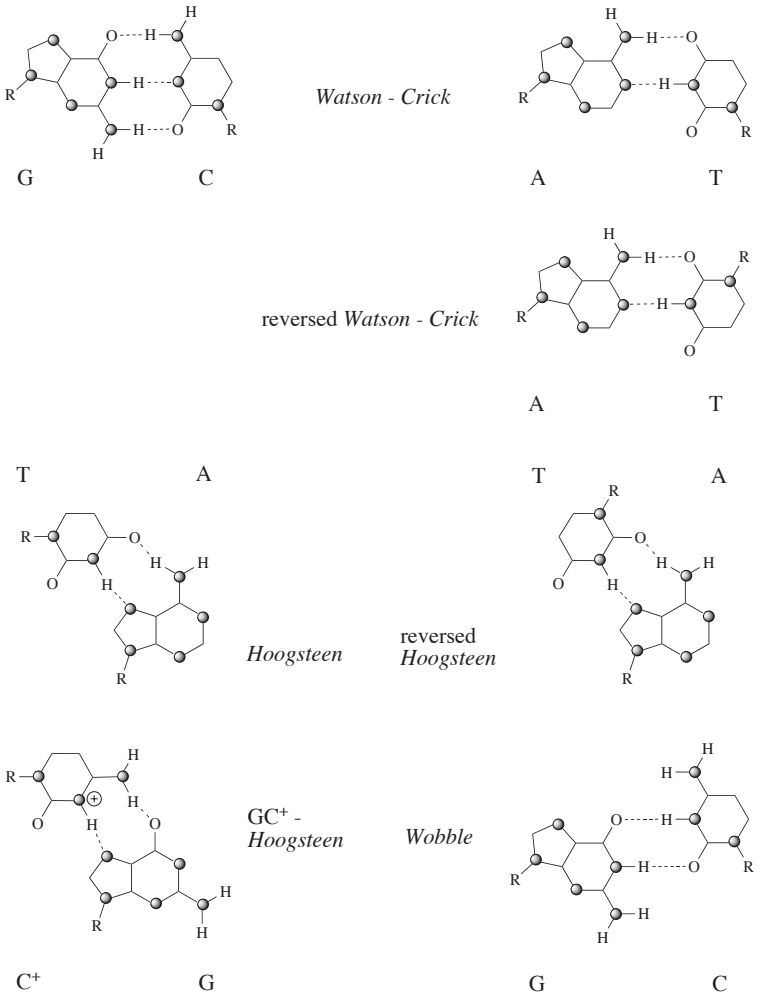


Figure 3.7: Simplified *non-Watson-Crick* bonding motifs.

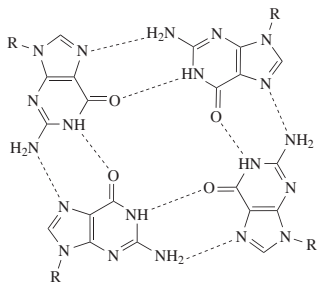


Figure 3.8: G-Quadruplex motif.

3.3.2.7 Stability of Nucleobases, Nucleosides and Nucleotides

In terms of stability, there are three parts of the DNA structure that are susceptible to spontaneous decomposition or cleavage in the cellular environment: the nucleobase, the glycosidic bond and the phosphodiester. Especially in terms of chemical evolution, the stability of present-day nucleobases is of interest. High temperature conditions for the origin of life is widely favored among experts because hyperthermophiles, which are claimed to be the oldest organisms on earth, grow at temperatures between 80 °C and 110 °C.[6,126–128]

M. Levy and *S.L. Miller* investigated the stability of nucleobases at various temperatures and collected the following data (*Figure 3.9, Table 3.3*).[129]

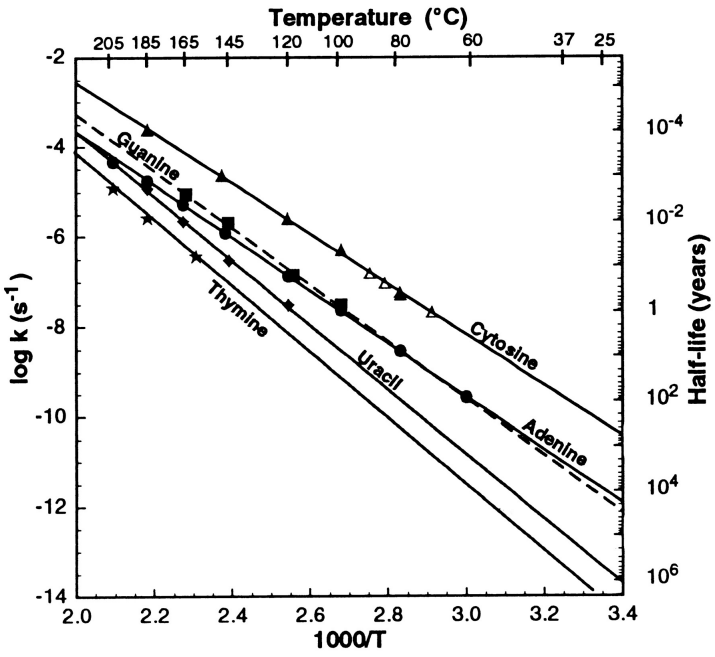


Figure 3.9: Arrhenius plot for decomposition of A, U, G, C and T (pH = 7).

The most frequent kind of decomposition of nucleobases is deamination. Cytosine can be converted to uracil by hydrolysis, adenine is converted to hypoxanthine and guanine to xanthine (see *Scheme 3.10*).[130]

The deamination of A and G is approximately 50-fold slower than those of C,[131] but, in view of absolute rates, the environment plays a crucial role. At 37 °C, the half-life ($t_{1/2}$) of cytidine in single-stranded DNA is ca. 200 years, which is quite similar to free cytosine, whereas in double-stranded DNA, it is

Table 3.3: Half-lives of nucleobases in *aq.* solution (conc. 10^{-3} M, 0.05 M phosphate buffer, pH = 7).

Temperature	A	G	U	T	C
100 °C	1 yr	0.8 yr	12 yr	56yr	19 days
25 °C	1×10^4 yr	1×10^4 yr	n.a.	n.a.	340 yr
0 °C	6×10^5 yr	1.3×10^6 yr	3.8×10^8 yr	20×10^8 yr	1.7×10^4 yr

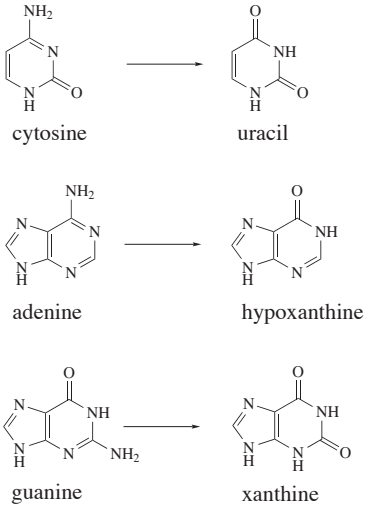


Figure 3.10: Deamination of C, A and G.

on the order of 30,000 years. This observation means that the deamination of C is 150-fold slower in duplex DNA and therefore strongly dependent upon tertiary structure.[132]

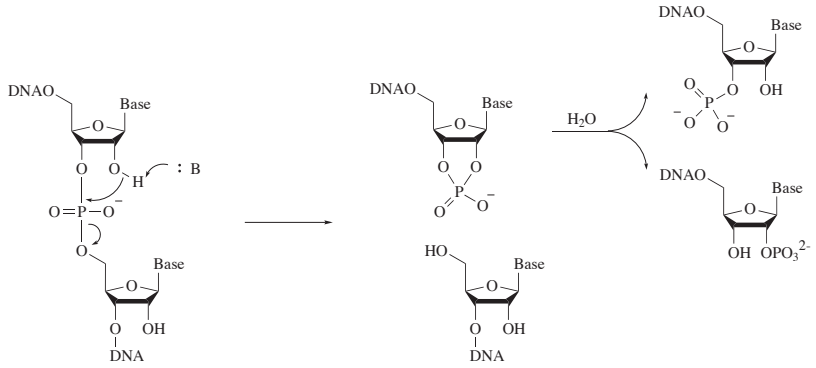
Deamination plays an important role in mutagenesis. For example, the hydrolytic deamination product of adenine, hypoxanthine, forms a more stable base pair with C than with T and needs to be removed by DNA repair mechanisms. The same is necessary for the cytosine hydrolysis product. The deamination of C plays a central role in this thesis and will be discussed in details in *section 3.4.1.* A remarkable fact should be mentioned here to show, how a marginal structural change can influence the stability of a nucleobase

and its biochemical relevance: 5-methylcytosine, a naturally occuring derivative of C, which constitutes 3% of DNA cytosine residues in mammalian cells, is deaminated three to four times more rapidly than cytosine. This relatively small difference is by itself unremarkable, but an inefficient DNA repair mechanism results in a dangerous mutagenic lesion.[133]

In addition to deamination, a few other endogenous damages are known. Oxidation plays an important role in mutagenic lesions. The prevalent damages are oxidation of guanosine to 8-oxoguanosine and Fapy·dG (an imidazole ring-opening product, which will be explained in details in *section 3.4*).
Another cause for DNA damage in living cells is nonenzymatic DNA methylation, for instance. *S*-adenosylmethionine (SAM; a cofactor in enzymatic transmethylation reactions) or other methyl donor agents are present in the cell nucleus and act as endogenous genotoxic agents.[134]

Regarding the stability of nucleosides and (oligo-)nucleotides, the *N*-glycosidic bond is a vulnerable point of attack. The hydrolytic release of free base was found to be a function of temperature, pH, ionic strength and nucleic acid secondary structure.[135,136] Generalized purine nucleotides were found to liberate the nucleobase 20-fold more rapidly than pyrimidine nucleotides. The velocity of depurination differs in single-stranded and double-stranded DNA only fourfold, Which means that the helical structure doesn't provide much protection. The influence of ionic strength was described to be marginal.

The difference between DNA and RNA, due to only one hydroxy group, is tremendous. On one hand the 2'-OH group stabilizes the *N*-glycosidic bond in RNA compared to DNA, but, on the other hand, it makes the phospodiester much more susceptible. The RNA phosphodiester bond is very frail to hydrolysis, particularly in the presence of base and/or divalent cations (Mg^{2+}, Ca^{2+}, Zn^{2+}).[137]

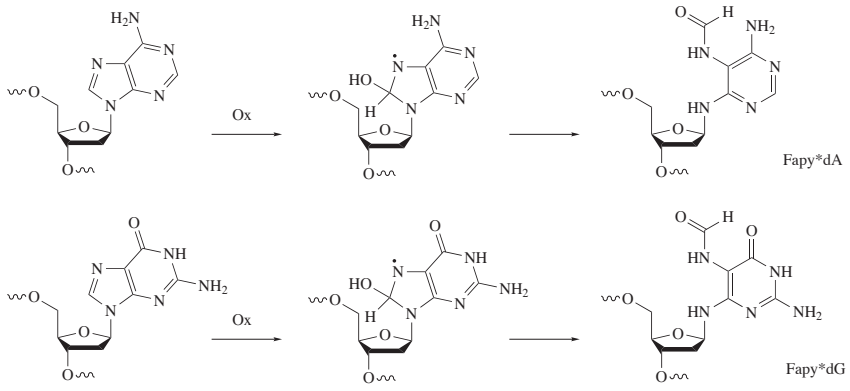


Scheme 3.1: Hydrolysis of RNA.

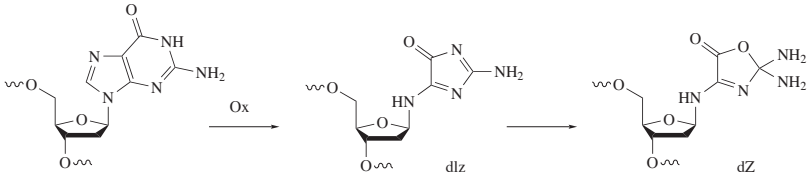
3.4 Exocyclic Amino Nucleosides (EANs)

Exocyclic amino nucleosides (EANs) are known in nature as well as in synthesis, but there is only a marginal knowledge about their biological significance and activity as drugs. EANs were found as decomposition products of purine nucleosides, as a fungal metabolite, as well as intermediates in the degradation metabolism of guanosine.[61,138,139] A few examples were investigated as antitumor and antiviral drug candidates due to their ability to mimic purine nucleobases.[140,141]

Fapy·dA and Fapy·dG are degradation products of dA and dG, for instance. These formamidopyrimidine lesions are produced in DNA as a result of oxidative stress and have been implicated in mutagenesis.[142–147] Purine nucleotides, dG especially due to their low oxidation potential, are sensitive to oxidative stress because a formal addition of a hydroxyl radical (or singlet oxygen $^1O_2$) can lead to opening of the imidazole ring (*Scheme 3.2*). The specific details of the mechanisms for the formation of these lesions are uncertain.[64] It has also been shown that oxidative degradation of dG can lead to two other EANs: dZ and dlz. Although the mechanisms of these lesions are not certain yet, the damage seems to be a consequence of oxidation of dG to a guanine radical cation ($dG^{\bullet+}$) resulting in 2-aminoimidazolone (dlz) and its hydrolysis product, oxazolone (dZ) as seen in *Scheme 3.3*.[138]



Scheme 3.2: Formation of Fapy·dA/dG.



Scheme 3.3: Formation of dG lesions dlz and dZ.

EANs can also be found in natural product biosynthesis. Clitocine, a secondary fungal metabolite, is probably the most investigated EAN and was isolated from the mushroom *Clitocybe inverse*.[61,148] Clitocine (6-amino-5-nitro-4-(β-D-ribofuranosylamino)pyrimidine) is an inhibitor of adenosine kinase and shows strong insecticidal activity and an intense cytostatic effect towards several leukemia cell lines.[149]

Another pyrimidine EAN found in nature is 2-(hydroxy/amino)-6-hydroxy-5-(*N*-methyl-*N*-formylamino)-4-(D/L-ribofuranosylamino)pyrimidine.[139] This EAN seems to be a metabolic intermediate of guanosine degradation toward derivatives of folic acid (see *Figure 3.11*).

Only a few artificial EANs are reported in the literature, such as exocyclic amino pyrimido[5,4-*d*]pyrimidines[140,141] and triazine nucleosides.[63] The 4-substituted 8-(Deoxy-D-ribofuranosylamino)pyrimido[5,4-*d*]pyrimidines show antitumor and antiviral properies (*Fig. 3.12*).

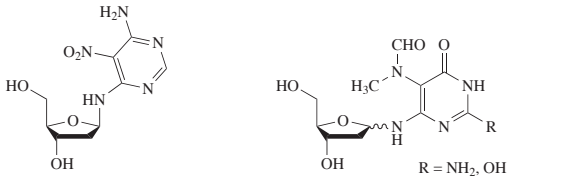


Figure 3.11: Clitocine and a degradation product of G.

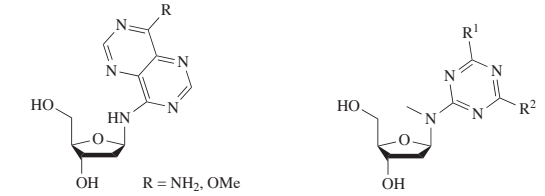
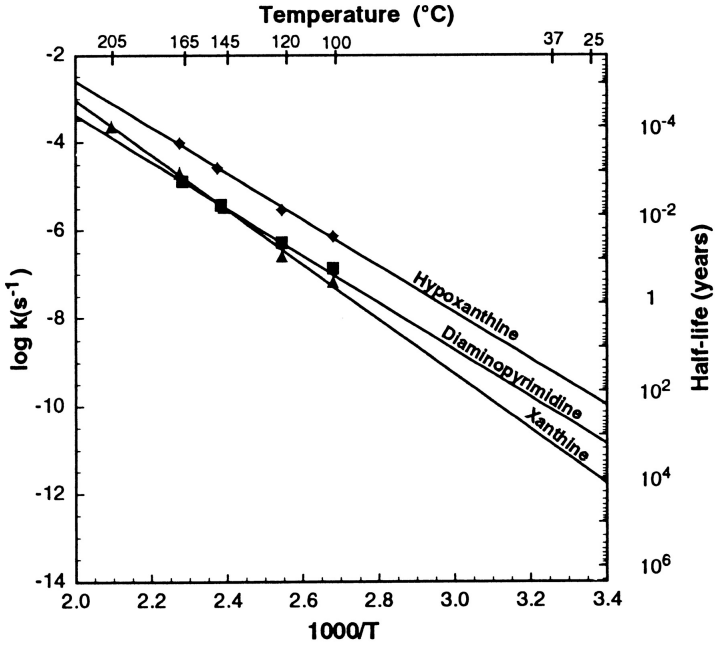


Figure 3.12: Synthetic EANs.

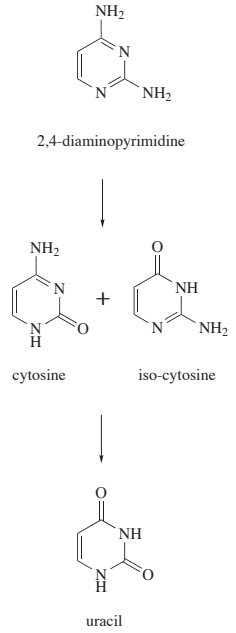
3.4.1 Hydrolysis of 2,4-Diaminopyrimidine and Cytosine

The fundamental idea behind this project is based upon the fact that 2,4-diaminopyrimidine (D) can be converted to cytosine and further to uracil by hydrolysis (see *Fig. 3.13b*). 2,4-diaminopyrimidine was presumable available on prebiotic earth and could act as source of cytosine and uracil. Possible syntheses of 2,4-diaminopyrimidine under prebiotic conditions were reported by *Miller* from guanidine hydrochloride and cyanoacetaldehyde[150] and by *Shapiro* from cyanat and cyanoacethylene.[151] *Levy* and *Miller* investigated the hydrolysis of D in detail.[129] As expected, they observed a mixture of cytosine and iso-cytosine from the first deamination step as well as fully hydrolyzed uracil. The half-life of D was found to be $t_{1/2}$ = 42 days at 100 °C and 40,000 yr at 0 °C. These hydrolysis rates are about the same as those of cytosine ($t_{1/2}$ = 19 days at 100 °C, $t_{1/2} \geq 10^6$ yr at 0 °C) and are fast on a

geologic time scale. This fact indicates that D and C would be excluded from use in a high-temperature origin of life. Additionally, other factors could affect the rates of hydrolysis, such as the presence of other compounds in the prebiotic soup. The deamination of cytosine to uracil, for example, has been shown to be catalyzed by sulfite[152] and other nucleophiles may act similarly. Summarized, these facts point to a low-temperature origin of life.



(a) Arrhenius plot for the decomposition of D (pH = 7)



(b) Hydrolysis of D

Figure 3.13: Diamination of 2,4-diaminopyrimidine.

3.4.2 Reactivity of EANs

The inertness of EANs was investigated in the case of formamidopyrimidines[64,153] and exocyclic amino triazine nucleotides (EATNs).[63] As seen in *Scheme 3.14*, EANs can undergo acid-induced anomerisation, isomerisation to the pyranose form, and hydrolysis (deglycosylation).

Greenberg et al. investigated Fapy·dA/dG lesions concerning epimerisation and deglycosylation, respectively.[64] For a monomeric Fapy·dA nucleoside, they observed anomerisation of an α/β 1:1 mixture in phosphate buffer (10

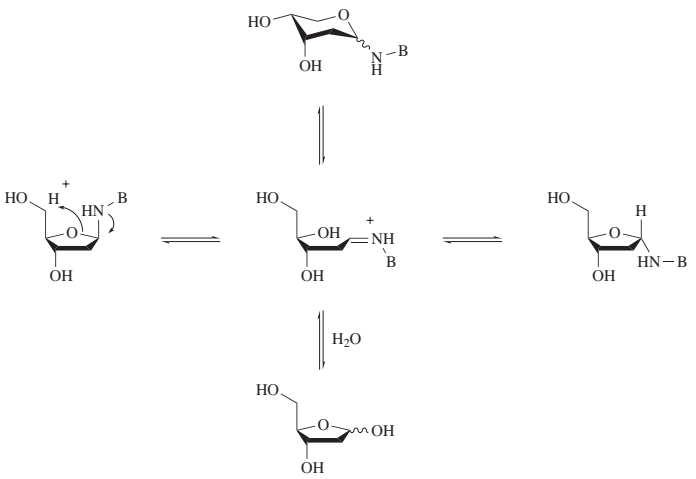


Figure 3.14: Anomerisation, isomerisation and hydrolysis of EANs.

mM, pH = 7.5, rt) shifting the ratio to α/β 1.33:1 within 6 h. Deglycosylation occurred in the same buffer system as well. Temperature dependent deglycosylation was determined for Fapy·dA and Fapy·dG in the monomeric form as well as incorporated into single stranded DNA (ssDNA). Cleavage rates under alkaline conditions (0.1 M NaOH) for monomeric nucleoside and ssDNA were determined to be in the same range. Fapy·dG has a half-life in the range of days, whereas Fapy·dA has a half-life in the range of weeks (*Table 3.4*). Fapy·dG is less reactive than Fapy·dA, possibly due to resonance stabilization (*Scheme 3.15*).

Table 3.4: Half-lives of Fapy·dA/dG.

Fapy·dA		Fapy·dG	
T [°C]	$t_{1/2}$ [h]	T [°C]	$t_{1/2}$ [h]
37	103		
55	20.5	55	514
72	11.3		
91	5.7	90	91

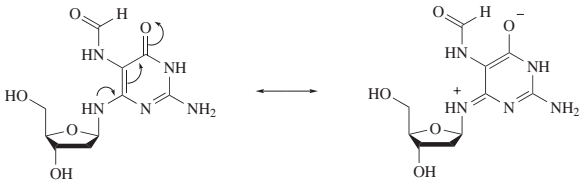


Figure 3.15: Resonance stabilization of Fapy·dG.

Also, *Burgdorf* and *Carell* performed stability studies on Fapy·dG by investigating the tetradiisopropyldisiloxane protected derivative (see *Figure 3.16*).[153] For the α-anomer, they found neither anomerization nor deglycosylation within 6 h in DMSO even at 60 °C. Furthermore, no decomposition occurred until the sample was heated to 100 °C for 24 h in DMSO and as expected (see *Section 3.3.2.7*) the ribose analogue was even more robust.

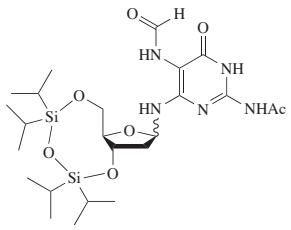
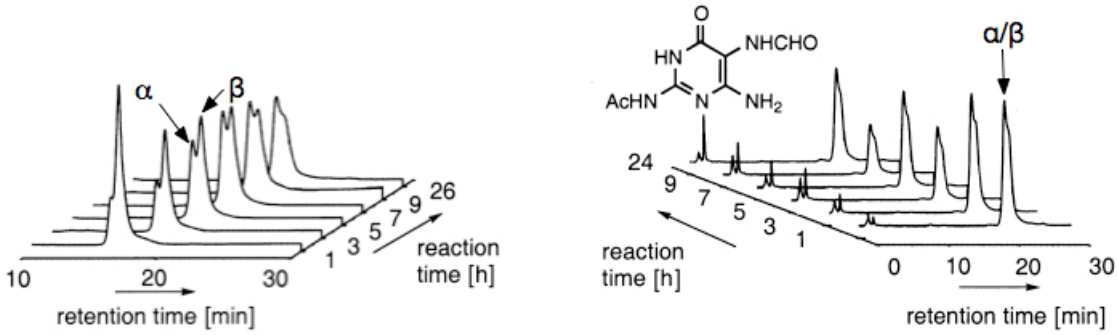


Figure 3.16: Disiloxane derivative of Fapy·dG.

In a more polar solvent (H_2O/MeCN 1:1), the β-anomer showed slow anomerization at room temperature (*Figure 3.17a*), but still almost no deglycosylation. The latter was observed to occur slowly at 50 °C as seen in *Figure 3.17b* (H_2O/MeOH 1:1, α/β-mixture 1:1). The half-lives were determined to be $t_{1/2}$ = 65.2 h for the α- and $t_{1/2}$ = 37.8 h for β-anomer.

A recent study on EATNs by *Hysell* and *Siegel* displayed the dependency of their stability on the electron deficiency of the nucleobase.[63] Various triazine derivatives were tested, in various solvents and pH. Electron-poor EATNs were quite stable under physiological conditions whereas electron-rich triazines anomerized, isomerized and deglycosylized quickly in diverse solvents

and at a broad range of pH. EATNs with intermediate electron density isomerized only in protic solvents and were stable at high pH.



(a) Anomerization of Fapy·dG *(b)* Deglycosylation of Fapy·dG

Figure 3.17: Stability of disiloxane-protected Fapy·dG.

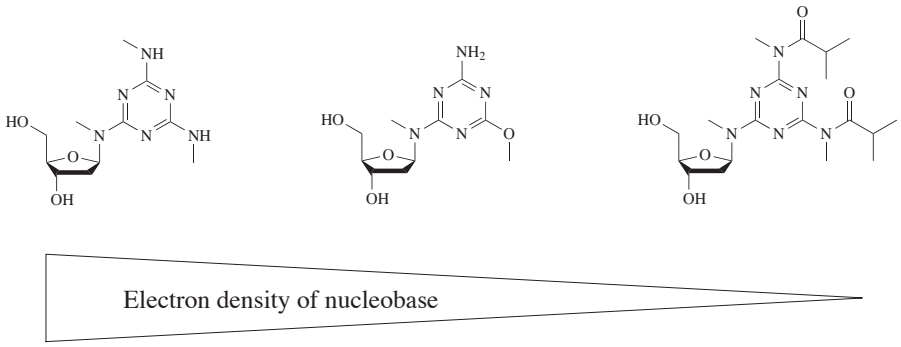


Figure 3.18: Electron density of investigated EATNs.

Degradation of the β-anomer of 4-amino-6-methoxytriazine EATN, which has an intermediate electron density, was followed by ^{1}H-NMR of the anomeric protons at (a) $t = 0$, (b) $t = 1$h, (c) $t = 9$h, (d) $t = 30$h (D_2O, pH = 6.0, *Figure 3.19*). One can see that anomerization occurs faster than isomerization to the pyranose. Nevertheless, the pyranose isomers (P-α, P-β) represent, as it was already reported for Fapy·dA[154] and Fapy·dG,[155] the thermodynamically favored products. Furthermore, a small amount of 2-deoxy-D-ribose was detected, indicating slow cleavage of the glycosidic bond relative to anomerization.

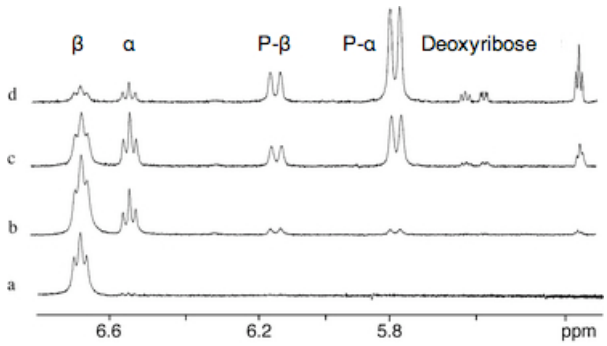


Figure 3.19: Anomerization, isomerization and deglycosylation of an EATN.

3.4.3 Usability of *C*-Analogues as Model Compounds

C-Nucleoside is a widely use term, but originally constituted nucleosides containing a C–C glycosidic bond instead of the naturally occurring C–N connection between the sugar moiety and the heterocyclic nucleobase.[156] In nature, a few *C*-Nucleosides are known (see *Figure 3.20*).[157,158] They often act as antibiotics and many also exhibit anticancer and/or antiviral activity. More recently, the term "*C*-Nucleoside" was used in an expanded fashion for several kinds of modifications like expanded *C*-glycosidic bonds or *N*→*C* substitutions of nucleobase heteroatoms. Some examples are shown in *Figure 3.21*.[159–162]

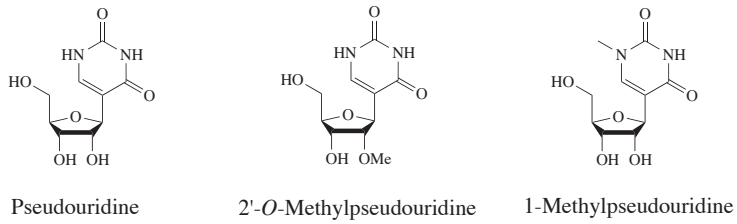


Figure 3.20: Naturally occurring *C*-nucleosides.

Another class of modern *C*-nucleosides are homo-*C*-nucleosides, which were first introduced by *Gensler* et al.[163,164] following many others.[165–167] Homo-*C*-nucleosides feature a methylene bridge between the anomeric carbon atom of the sugar portion and a carbon in the aglycon (nucleobase).

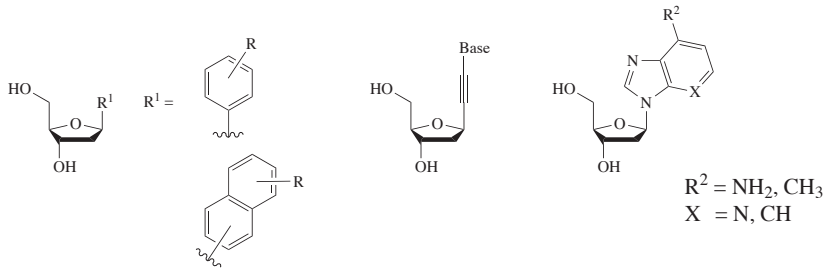


Figure 3.21: Examples of artificial *C*-nucleosides.

Greenberg et al. used a methodology developed by *Townsend* et al.[168] to synthesize homo-*C* analogues of Fapy·dA and Fapy·dG in order to avoid lability problems, such as anomerization, isomerization, and deglycosylation (see *Section 3.4.2*).[169] He showed computationally by molecular modeling studies and experimentally by duplex melting studies that the homo-*C* anologs are good models for the formamidopyrimidine lesion.[65,170]

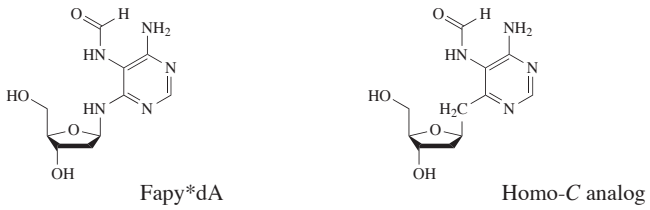


Figure 3.22: Fapy·dA and its homo-*C* analog.

Additionally, *Berstis* showed computationally in her master thesis that the electron distributions in nucleobases 2-aminopyrimidine (D)[1] and 2,4-diaminopyrimidine (D-NH2) are quite similar to their natural analogue adenine.[66] The calculated electron distributions of the single nucleobases D (nucleobase of the corresponding homo-*C* nucleoside), and D-NH2 (nucleobase of the corresponding EAN) as well as pairing of those with uracil show clear similarity to adenine and its pairing to uracil (see *Figure 3.23*).

[1]Nomenclature was adopted from *Berstis*, but is systematically incorrect.

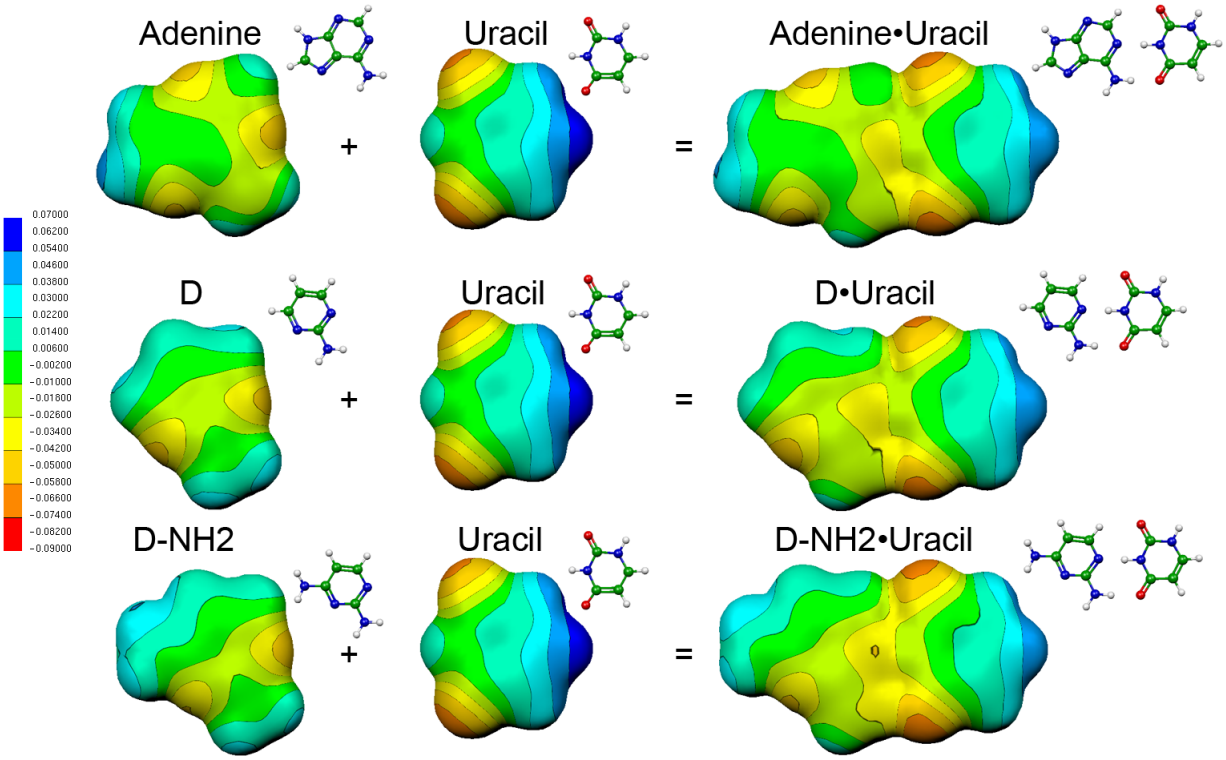


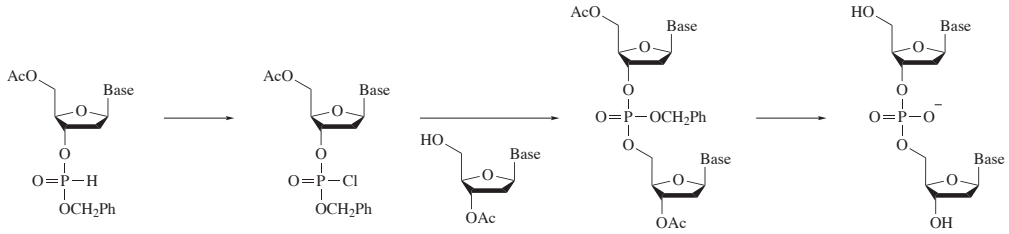
Figure 3.23: Calculated electron distributions of nucleobases A and D.

Based on the cognition that homo-*C* nucleosides seem to be good models for EANs, *Löpfe* synthesized a homo-*C* pyranose nucleic acid as a model of the corresponding pyranose EAN.[171] He was able to investigate its base pairing ability, which is in structural aspects similar to its EAN analogue.

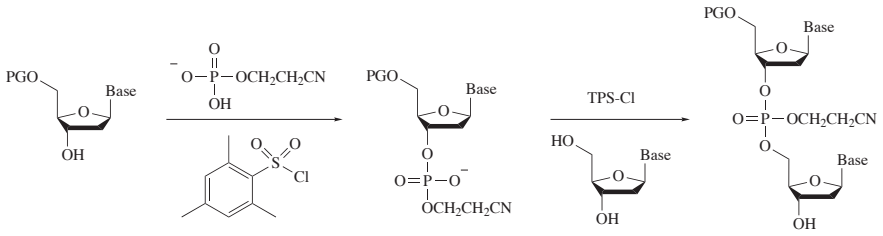
Figure 3.24: Pyranose EAN and its homo-*C* analog.

3.5 Synthesis of Oligonucleotides by the Phosphoramidite Approach

The discovery and elucidation of the DNA structure provided the motivation to find a way to synthesize oligonucletoides of defined sequences.[72,81] *Michelson* and *Todd* reported the first chemical synthesis of a dinucleotide in 1955 following the synthetic approach seen in *Scheme 3.4*.[172] Through the years, many approaches have been developed and modified by introducing new phosphorylating and condensing reagents but at least one milestone should be mentioned here. *Letsinger* et al. have developed the so called "phosphite triester" methodology (*Scheme 3.5*).[173,174]



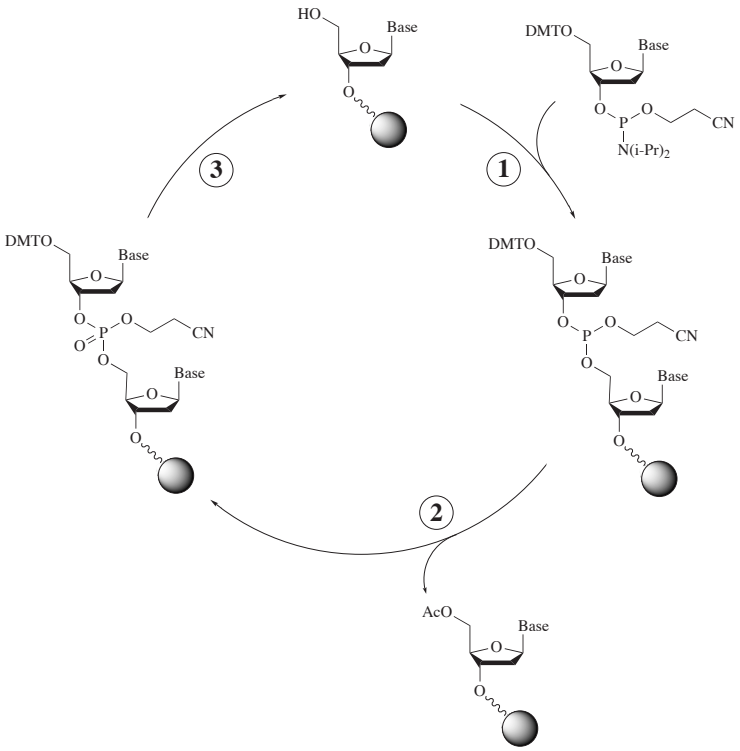
Scheme 3.4: Dinucleotide synthesis.



Scheme 3.5: "Phosphite triester" methodology.

The approach that is currently the most commonly used for oligonucleotide synthesis was evolved by *Beaucage* et al. and is only a small variation of the phosphite triester methodology: the *phosphoramidite approach*.[175] Enhancements of this methodology have led to an incredible collection of reagents and

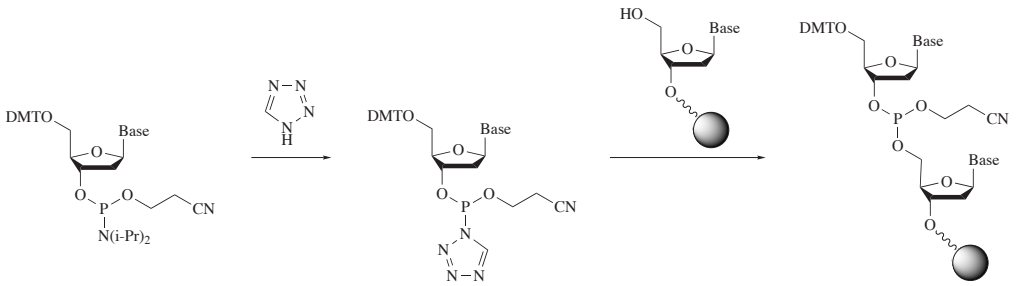
conditions, which allows fine tuning for countless specific problems. A broad review article by *Beaucage* summarizes the progress of the last decades.[176]



Scheme 3.6: "Phosphoramidite methodology".

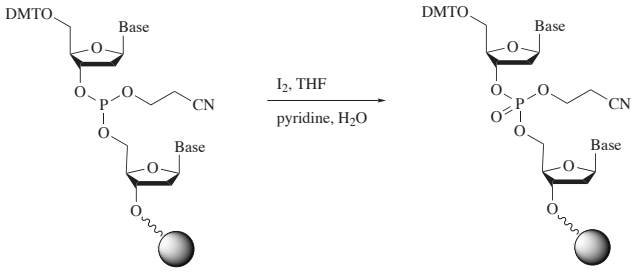
Oligonucleotide synthesis by the phosphoramidite approach is a 3'→5' chain elongation methodology. The reaction cycle contains four main steps: coupling, capping/oxidation, detritylation and a final cleavage of the support (*Scheme 3.6*).

1. *Coupling*: In the first step, activated (2-cyanoethyl)phosphoramidite (CE PA) reacts with a 5'-deprotected nucleotide bound to a solid support. Standard protocols for automated oligonucleotide synthesis use typically 1*H*-tetrazole for activation of CE PA, but, alternatively, other tetrazole derivatives, like 5-(ethylthio)-1*H*-tetrazole for instance, can be used (*Scheme 3.7*).



Scheme 3.7: Phosphoramidite activation.

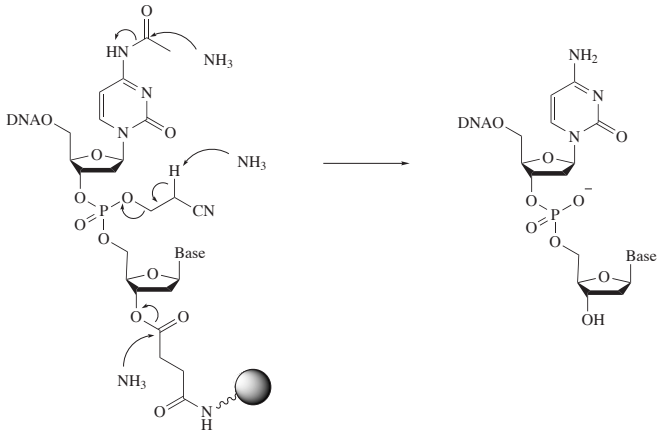
2. *Capping*: In order to protect uncoupled 5'-OH groups, a mixture of 1-methylimidazole in THF and 2,6-dimethylpyridine in Ac_2O is used. *Oxidation*: After capping, the solid support bound oligonucleotide is treated with I_2 in THF/pyridine/H_2O to oxidize the phosphite (*Scheme 3.8*).



Scheme 3.8: Phosphite oxidation.

3. *Detritylation*: In the last step of the cycle, the protection group (DMT) at the 5'-position is cleaved with trichloroacetic acid in CH_2Cl_2 to make the 5'-end available for the next coupling cycle.

4. *Clevage*: After finishing the synthesis, the oligonucleotide needs to be released from the support and fully deprotected. This is accomplished with *conc. aq.* NH_3-solution, which cleaves off all protection groups on the nucleobases and phosphates and additionally releases the linker to the solid support (*Scheme 3.9*).



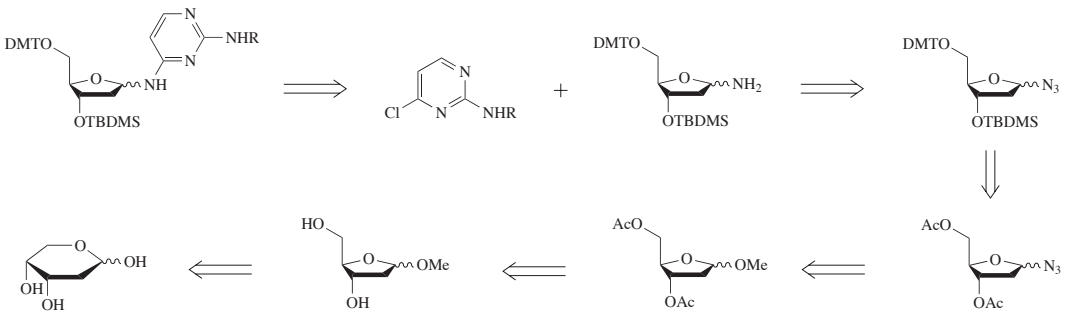
Scheme 3.9: Final cleavage.

Chapter 4

Own Work

4.1 Synthesis towards EAN

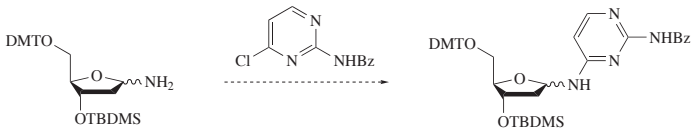
EANs can undergo anomerisation, furanose-pyranose isomerisation and cleavage of the glycosidic bond as known from work of *Hysell*,[63] *Greenberg*[64] and *Cadet*[154,155] (see *section 3.4.2*). In order to investigate the inertness of our two target exocyclic amino nucleosides D and E, we attempted to synthesize them following a strategy developed by *Greenberg* (*Scheme 4.1*).[169]



Scheme 4.1: Retrosynthetic analysis.

Greenberg developed this strategy for the synthesis of Fapy·dG, a DNA lesion caused by oxidative stress as seen in *section 3.4*. Fapy·dG contains an electron-poor pyrimidine and therefore differs significantly from our system. The key step of this synthesis is the final nucleophilic attack of a free amino

group at the chlorinated pyrimidine (*Scheme 4.2*), a reaction which seems to work only with electron-poor aromatic rings. The coupling failed in our hands under various conditions. Consequently *Greenberg's* methodology is not generally applicable. These synthetic problems and the expectation of an labile product persuaded us to develop a model system which is synthetically available, thermodynamically stable, and sterically and electronically comparable.



Scheme 4.2: Unsuccessful coupling.

4.2 Development of a Model System

Based on *Greenberg's* and *Berstis's* results, which showed carbon analogues to be useful models for EANs,[66,169] we decided to follow this approach. Homo-*C* nucleosides were not expected to be susceptible to either anomerisation/isomerisation or deglycosylation. They should be sterically quite similar, not dramatically different electronically, and therefore good models for the purpose of investigating configurationally stable analogs of EANs in oligonucleotides.

A model system based on a four letter code would therefore consist of the two natural deoxynucleosides dC and dU and, in addition, two artificial homo-*C* deoxynucleosides dD and dE (*Scheme 4.1*). This system allows for the study of base pairing abilities with stable species in the absence of an anomerisation and isomerisation equilibrium.

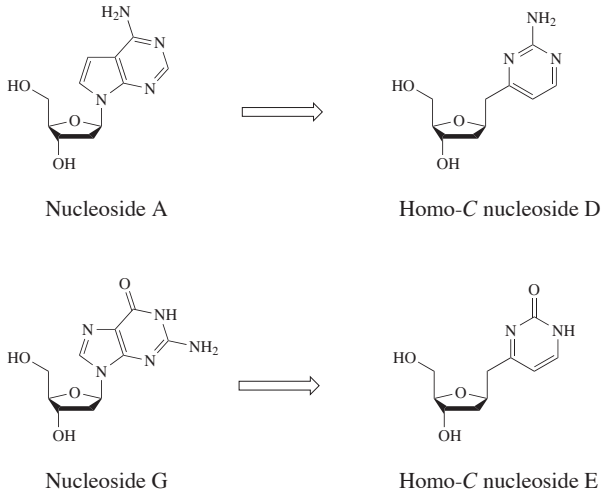


Figure 4.1: Homo-*C* nucleosides as analogues of A and G.

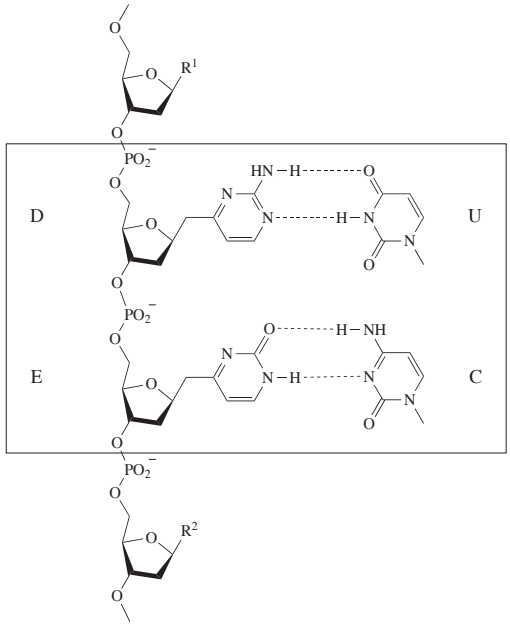
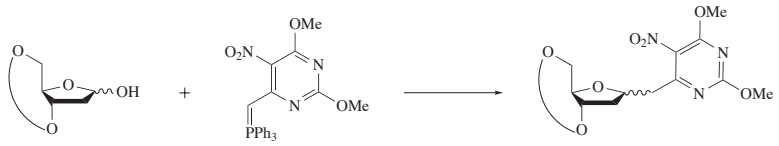


Figure 4.2: Model system with homo-*C* deoxynucleosides dD and dE.

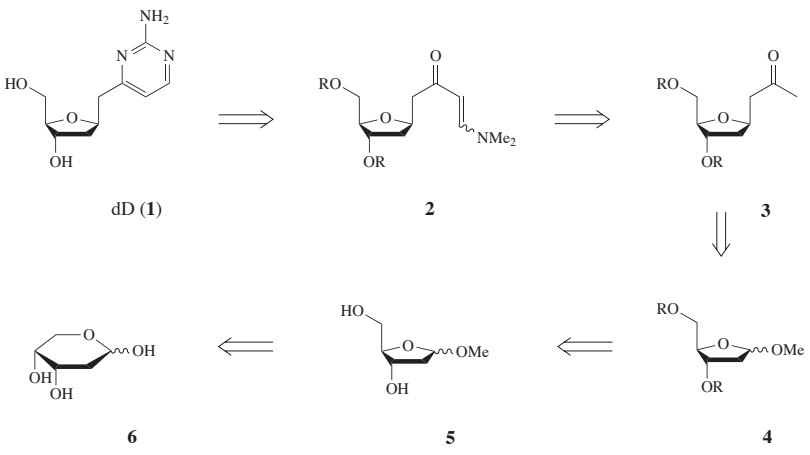
4.3 Retrosynthetic Analysis

Greenberg et al. applied homo-*C* nucleosides by forming the glycosidic bond with preconstructed pyrimidine derivatives *via* a *Wittig-type* reaction, a method first introduced by *Townsend* et al. (*Scheme 4.3*).[168]

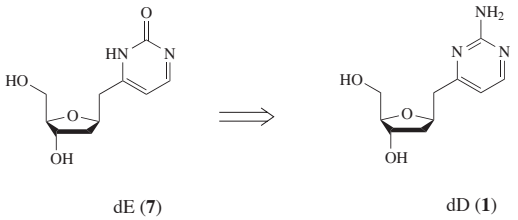


Scheme 4.3: Wittig reaction.

Due to the fact that a *Wittig* reaction is not compatible with an amino function, we decided to use another strategy and build up the pyrimidine ring by condensation. Nucleoside dD (**1**) was prepared from enaminoketone **2**, which acts as an equivalent of a β-ketoaldehyde, by condensation with guanidine (*Scheme 4.4*). Unfortunately nucleoside dE (**7**) can not be accessed in the same way, due to the poor nucleophilicity of urea, but **1** can be transformed to **7** *via* a diazonium ion (*Scheme 4.5*). Enaminoketone **2** can be synthesized with *Bredereck's* reagent[177] from ketone **3**, which is in turn available from acetal **4**, a simply reachable derivative of 2-deoxy-*D*-ribose (**6**). Because the synthetic results showed the methods used were not diastereoselective, the synthesis was carried out as an anomeric mixture. Separation of α- and β-anomers was achieved by HPLC at the final stage.



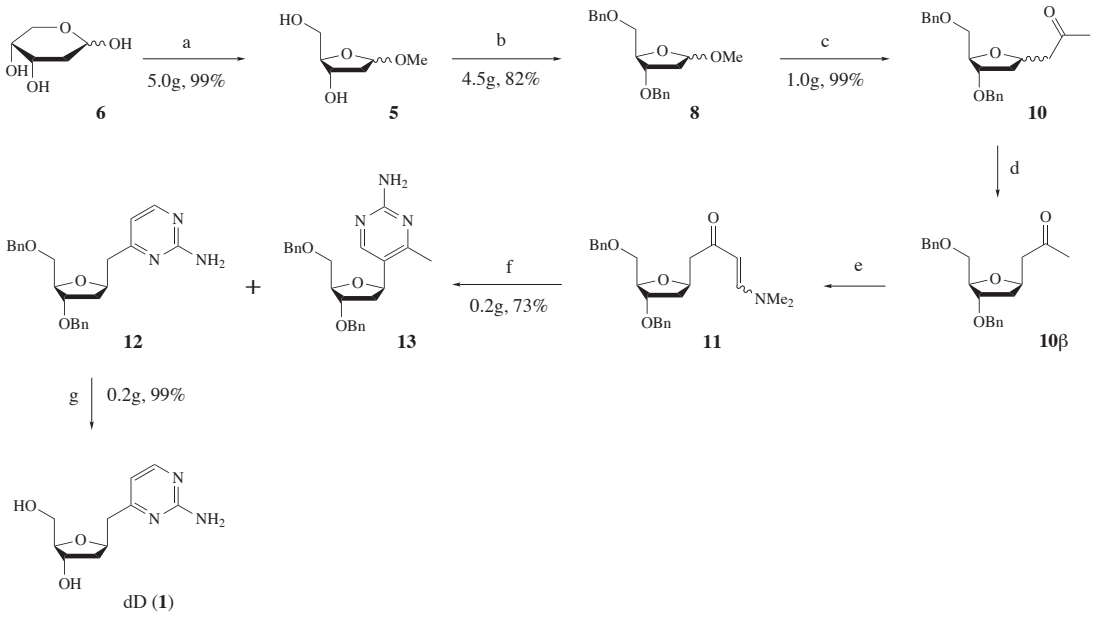
Scheme 4.4: Retrosynthetic analysis dD (**1**).



Scheme 4.5: Retrosynthetic analysis dE (**7**).

4.4 Synthesis of dD and its CE PA

Nucleoside dD (**1**) was synthesized in seven steps from 2-deoxy-*D*-ribose (**6**) (*Scheme 4.6*).

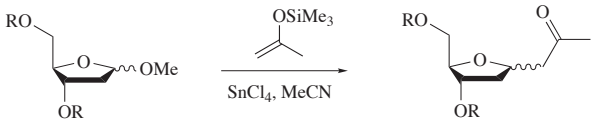


(a) i) MeOH, HCl, rt, 15min; ii) Ag_2CO_3. (b) i) NaH, THF, 0°C, ; ii) $INBu_4$, BnBr, rt, 4h. (c) $SnCl_4$, **9**, MeCN, 0°C, 1h. (d) $Zn(OAc)_2$, NaOMe, MeOH, rt, 24h. (e) $^tBuOCH(NMe_2)_2$, Tol, 70°C, 20h. (f) i) EtOH, NaOEt, 75°C, 5h; ii) $(CH_5N_3)_2{*}H_2SO_4$, 40°C to 70°C, 15h. (g) BBr_3, DCM, 0°C, 3min.

Scheme 4.6: Synthesis of dD (**1**).

According to the literature[178] **6** has been converted to acetal **5** almost quantitatively. Benzyl protection of the free hydroxy groups[179] gave **8** as an anomeric mixture of α/β ~1.2:1, which was separable by column chromatog-

raphy, but was used as a mixture for further synthesis. Coupling of **8** with trimethyl(prop-1-en-2-yloxy)silane (trimethylsilylenolether of acetone, **9**) in the presence of $SnCl_4$ formed the homo-*C* glycosidic bond (*Scheme 4.7*).[180–182] C(1) of the introduced propan-2-one unit represents essentially the methylene bridge of the later homo-*C* nucleoside.

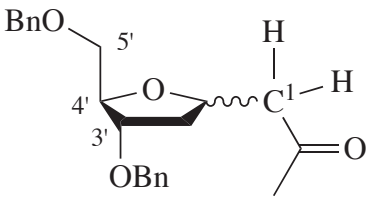


R = Bn: **8**
= Bz: **14**

R = Bn: **10**
= Bz: **15**

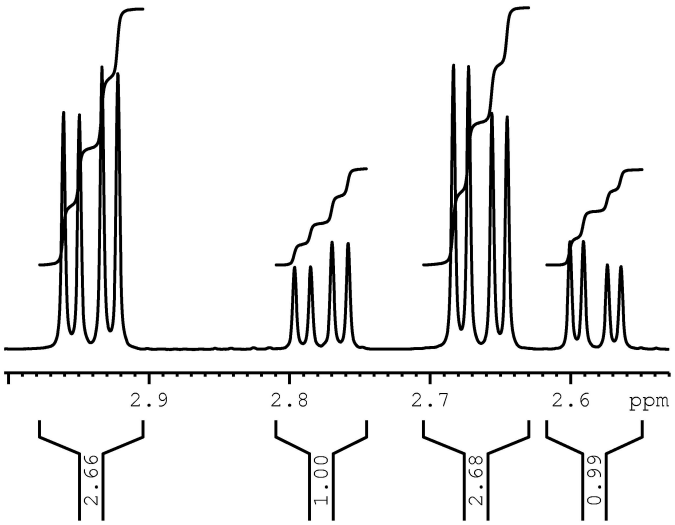
Scheme 4.7: $SnCl_4$ coupling.

This coupling worked in excellent yields if freshly distilled $SnCl_4$ was used. Unfortunately a separation of the anomers by column chromatography was not successful and determination of diastereomers and their ratios was performed by peak assignment with COSY and NOESY experiments and peak integration in ^{1}H-NMR (*Figures 4.3, 4.4*).



10

(a) Analyzed protons at C(1)



(b) ^{1}*H*-NMR of $H_2C(1)$ in ketone **10**

Figure 4.3: Determination of the anomeric ratio of ketone **10** by ^{1}H-NMR.

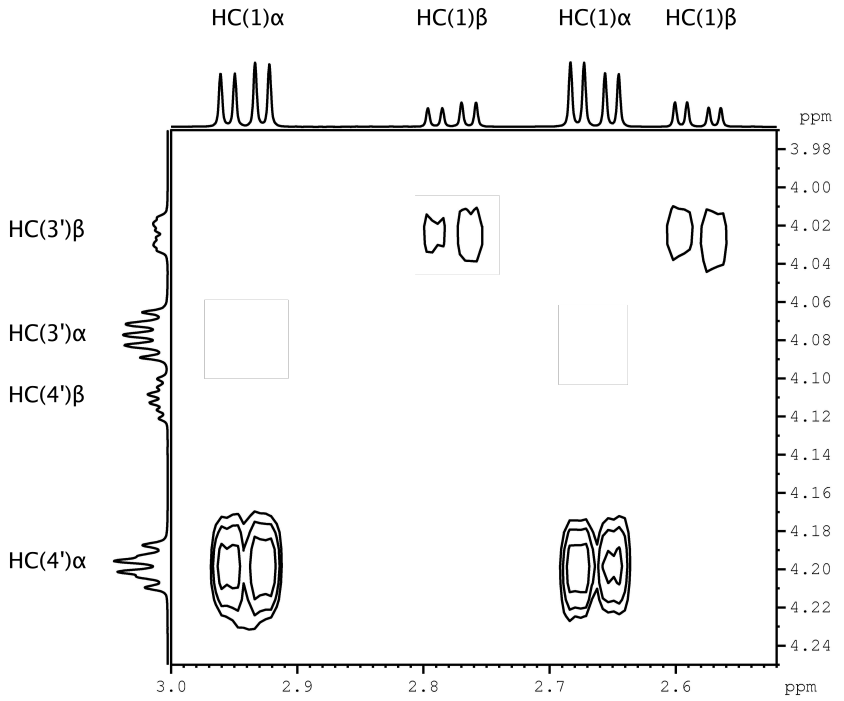
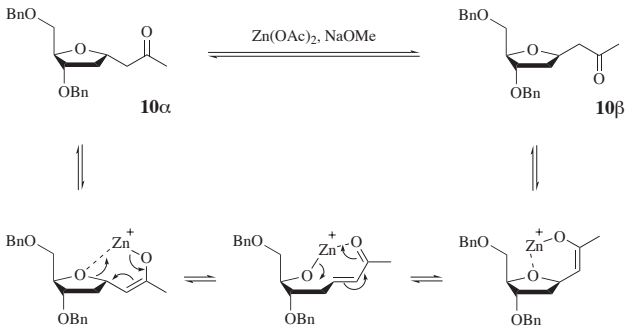


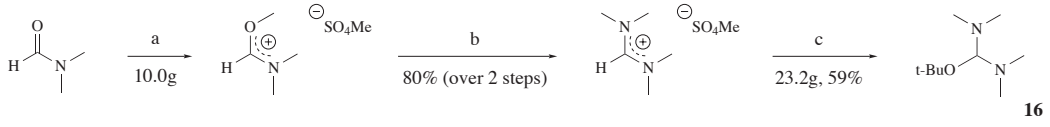
Figure 4.4: NOESY-NMR of ketone **10**.

The diastereoselectivity of this coupling reaction was found to be strongly dependent on the educt's configuration at C(1). If **8**α was used, **10** was obtained in a ratio α/β 5:1, but reaction with **8**β resulted in a ratio **10**α/**10**β 2.7:1. The same reactions were also carried out with benzoyl-protected acetal **14**, but the best result obtained was **15**α/**15**β 2:1. These results were unsatisfying since the α anomer is of marginal interest for this project. A Zn(II)-mediated epimerization developed by *Shao* et al. led us to an acceptable result.[183] Treating diverse anomeric mixtures of **10** with $Zn(OAc)_2$ under basic conditions shifted the ratio to a two-fold excess of β-anomer. *Shao* et al. hypothetized this epimerization to be initiated by the formation of Zn-enolate and it is known that Zn-enolates adopt only the *Z*-configuration,[184] which can be stabilized by intramolecular chelation to the furanose ring oxygen. This coordination weakens the C(1')–O bond and leads to a *retro Michael reaction*. Subsequent hetero-intramolecular *Michael addition* results in the formation of the β-homo-*C* glycoside (see *Scheme 4.8*).



Scheme 4.8: Epimerization of **10**.

In the next step of the synthesis, *Bredereck's* reagent was used to achieve enaminoketone **11**. Experiments performed by *Löpfe*[62] have shown *Bredereck's* reagent to be an effective derivative of DMF to access enaminoketone **11**, although a variety of formamide acetals are known to act similar.[185] *Bredereck's* reagent (*tert*-butoxybis(dimethylamino)methane, **16**) was synthesized from DMF as reported in the literature (see *Scheme 4.9*).[177,181]

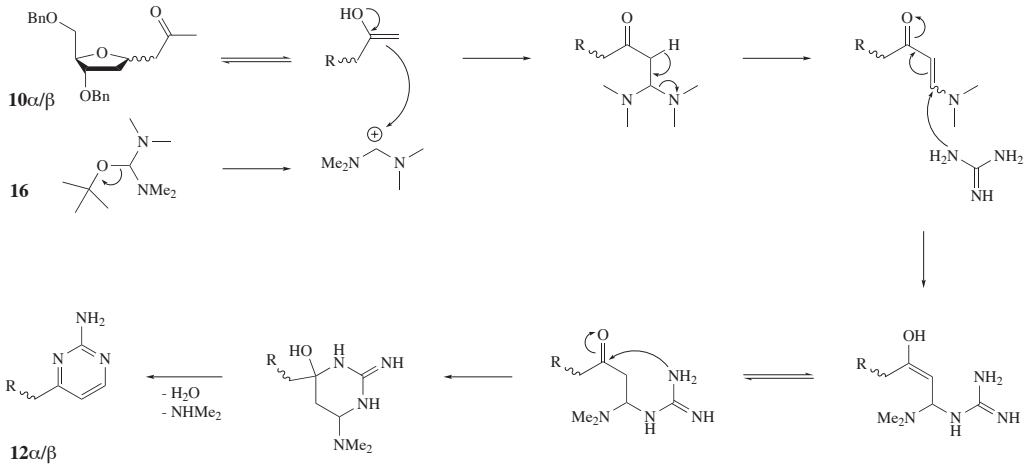


(a) i) Me_2SO_4, rt, 0.5h; ii) 60°C, 3h. (b) C_6H_6, $NH(CH_3)_2$, 80°C, 1h. (c) KO^tBu, rt, 1.5h.

Scheme 4.9: Synthesis of *Bredereck's* reagent (**16**).

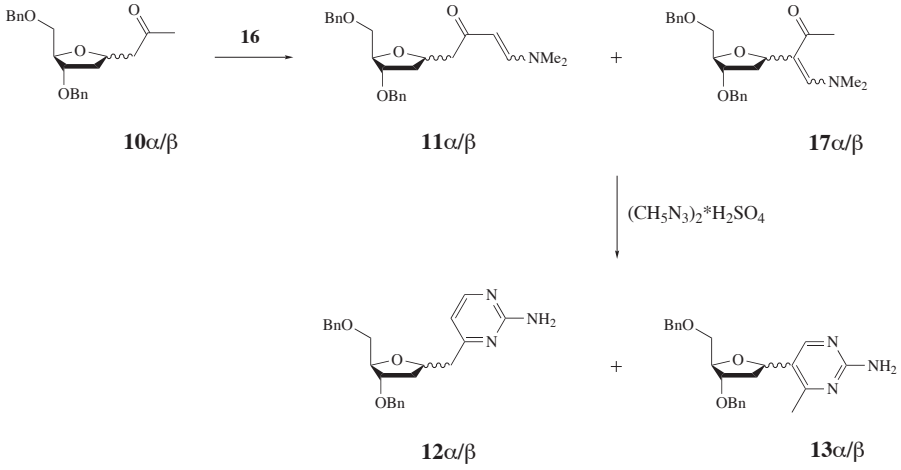
Compound **10** was converted with *Bredereck's* reagent (**16**) to enaminoketone **11**, which was not isolated, but treated with guanidinium sulfate under basic conditions in EtOH at 70 °C for 4 h, under which conditions a condensation to aminopyrimidine **12** takes place. A proposed mechanism for this reaction is shown in *Scheme 4.10*.

Unexpectedly, a mixture of isomers **12**α/β and **13**α/β was obtained in this reaction, which can be explained by an addition of *Bredereck's* reagent at the undesired position *C*(1) instead of *C*(3) (see *Scheme 4.11*). Under these reaction conditions a product ratio **12**/**13** 1:2 was observed. By screening



Scheme 4.10: Mechanism of guanidinium condensation.

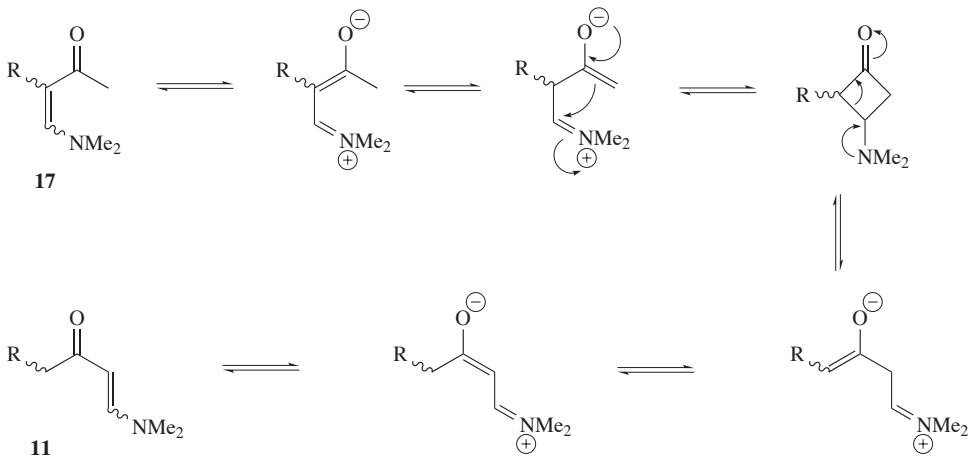
various reaction conditions, we have been able to shift this ratio toward the desired product (see *Table 4.1*). Heating the enaminoketone intermediate under basic conditions before adding guanidinium sulfate gave **12** as the main product, a fact that indicates an isomerisation taking place from **17**α/β to **11**α/β (see *Scheme 4.12*). This reaction needs to be controlled very carefully because a preheating of the enaminoketone under basic conditions over 60 °C leads to partial decomposition.



Scheme 4.11: Condensation to **12** and **13**.

Table 4.1: Tested reaction conditions for the pyrimidine condensation.

Reaction conditions	Ratio **12**/**13**	Yield [%]
EtOH, NaOEt, $(CH_5N_3)_2{*}H_2SO_4$, 70°C, 4h	1:2	73
i) EtOH, NaOEt, $(CH_5N_3)_2{*}H_2SO_4$, 40°C, 1.5h ii) EtOH, NaOEt, $(CH_5N_3)_2{*}H_2SO_4$, 70°C, 24h	1:1	56
i) EtOH, NaOEt, 60°C, 4h ii) EtOH, NaOEt, $(CH_5N_3)_2{*}H_2SO_4$, 70°C, 18h	6:1	58
i) EtOH, NaOEt, 70°C, 5h ii) EtOH, NaOEt, $(CH_5N_3)_2{*}H_2SO_4$, 70°C, 15h	85:1	13



Scheme 4.12: Proposed isomerisation.

Isolation of the four isomers **12**α/β and **13**α/β by column chromatography was not successful and necessitated a separation by HPLC, which turned out to be difficult as well. Finally, the separation was performed with a preparative *Spherisorb*®-NH_2 column and Hex/MeOH/EtOH 97:2:1 as eluant.

In *Figure 4.5*, one can see the chromatograms of two different isomeric mixtures. The upper chromatogram shows a mixture of **12**α/**12**β/**13**α/**13**β 1.7:1:1.9:3.3 resulting from a synthesis without an anomerisation step of ketone **10** and with unmodified guanidinium condensation[1]. The lower chromatogram displays a mixture of isomers obtained from an improved synthesis giving **12**β as the main product and **12**α, **13**α and **13**β in 30.0 %, 2.8 % and 9.2 % of the yield, respectively.

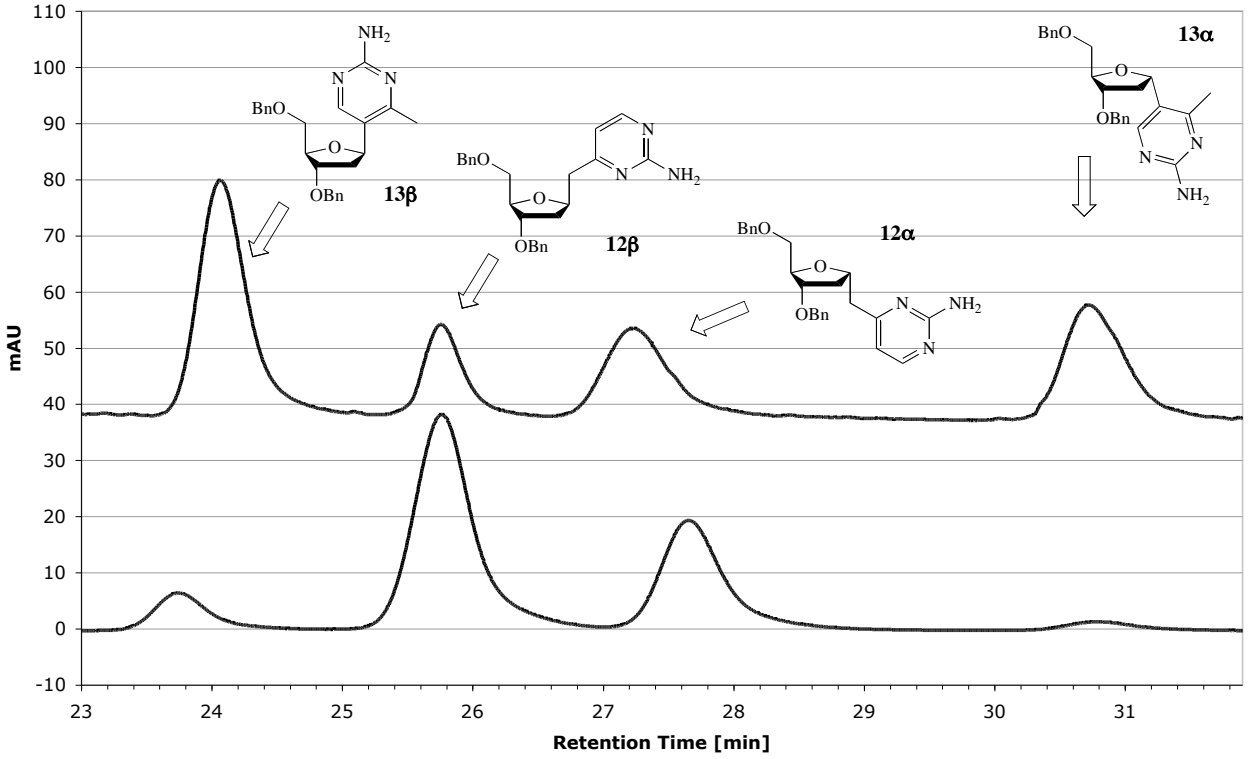


Figure 4.5: HPLC chromatograms of mixtures of **12**α/β and **13**α/β.

[1]Chromatogram is shifted up by 40 mAU for better clarity.

Identification of the isomers was achieved by NMR experiments. *Figures 4.6* and *4.7* show the ^{1}H-NMR of **12**β and the assignment of its peaks which was confirmed by *DQF-COSY*-, *NOESY*-, and *HSQC*-NMR. A sector of the *NOESY* spectra of **12**β is displayed in *Figure 4.8*, where one can see a *NOE* correlation between HC(1') and HC(4') but not between HC(1') and HC(3'). The reversed correlations can be found in **12**α as well as in **13**α, whose *NOESY*-NMR is shown in *Figure 4.9*. An additional proof that our classifications of the different isomers is correct is given by the *X-ray* structure we obtained from **13**α, unfortunately, the only isomer we have been able to crystallize (see *Figure 4.10*).

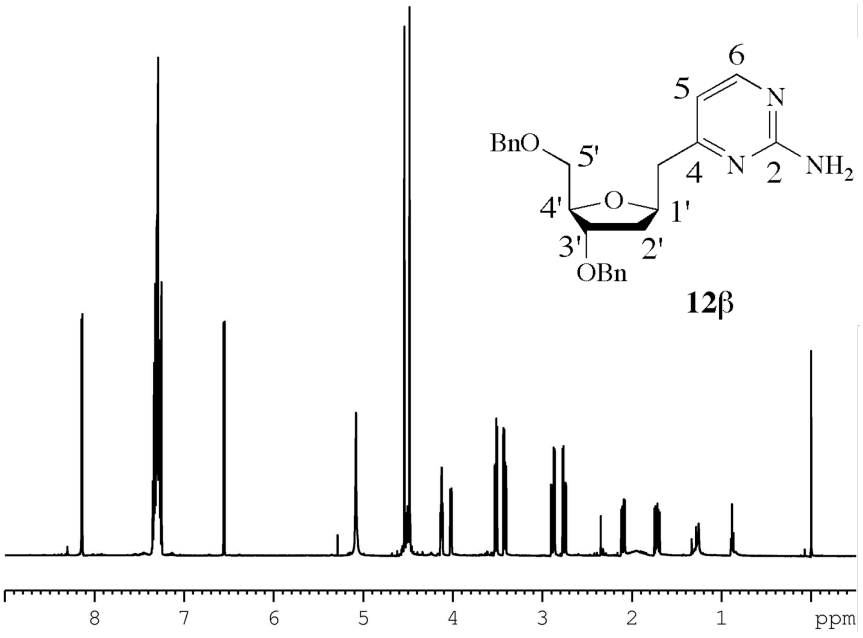
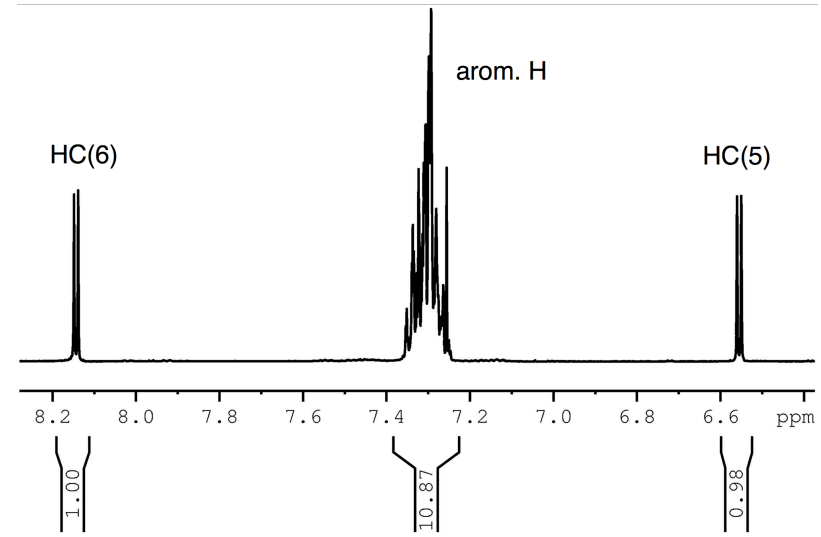
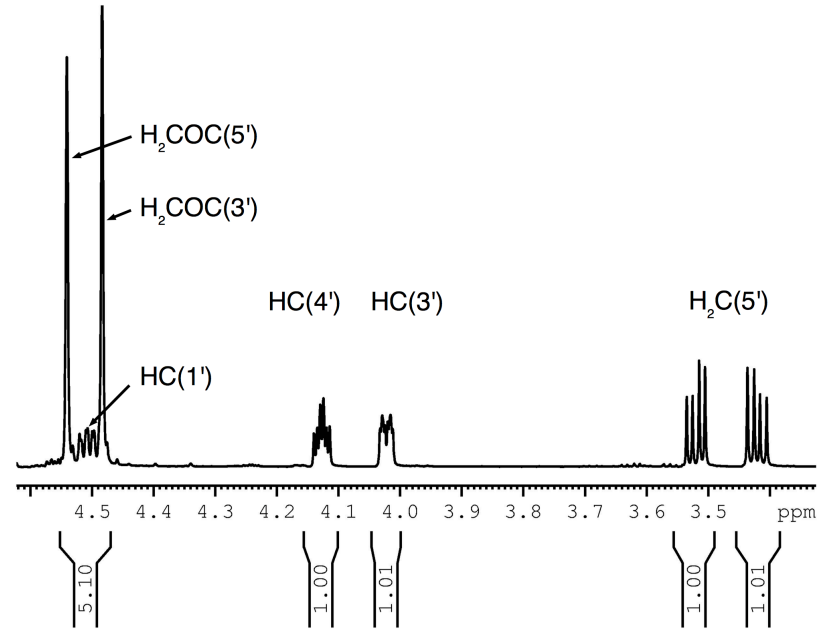


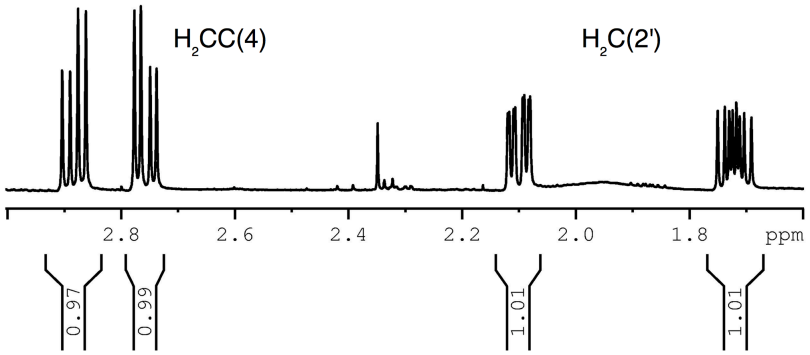
Figure 4.6: ^{1}H-NMR of **12**β.



(a) Aromatic region



(b) Benzylic protons, HC(1'), HC(3'), HC(4') and $H_2C(5')$



(c) Protons of methylene bridge and $H_2C(2')$

Figure 4.7: Detailed sections of ^{1}H-NMR of **12**β.

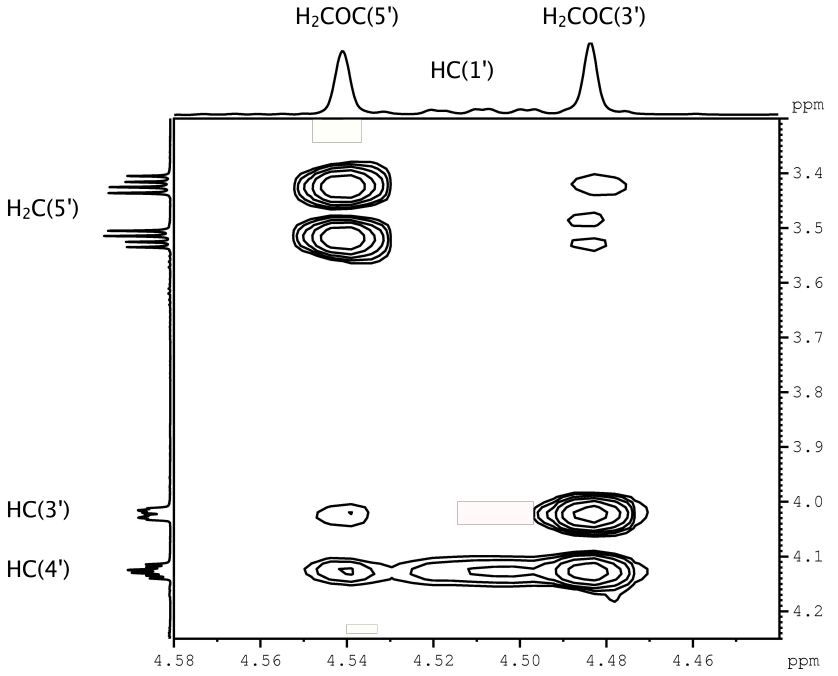


Figure 4.8: NOESY-NMR of **12**β.

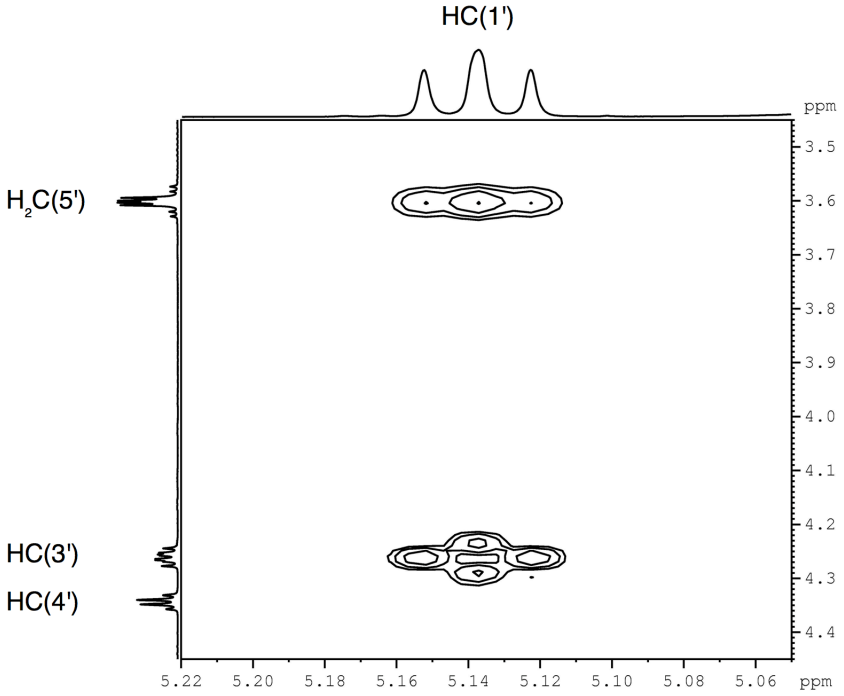


Figure 4.9: NOESY-NMR of **13**α.

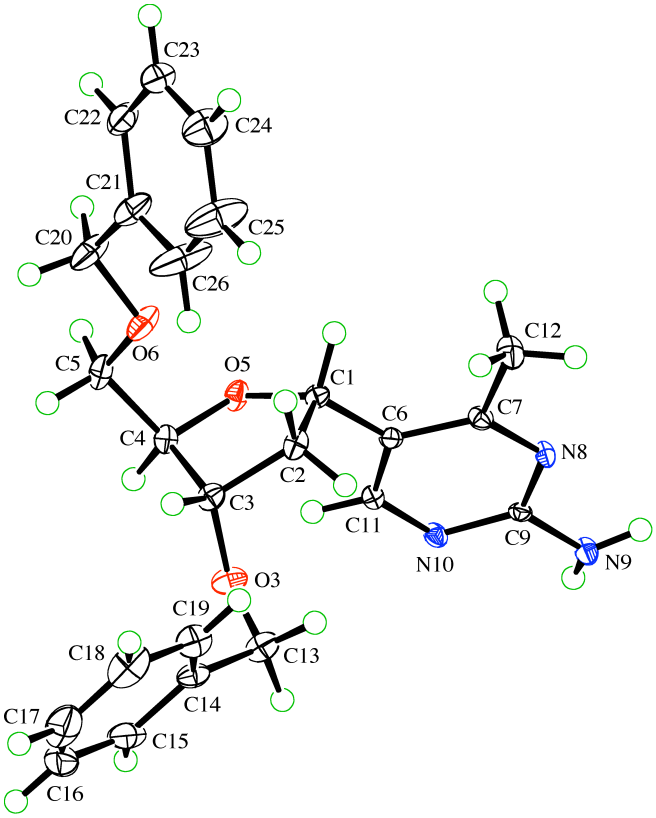


Figure 4.10: Solid state conformation of **13**α.

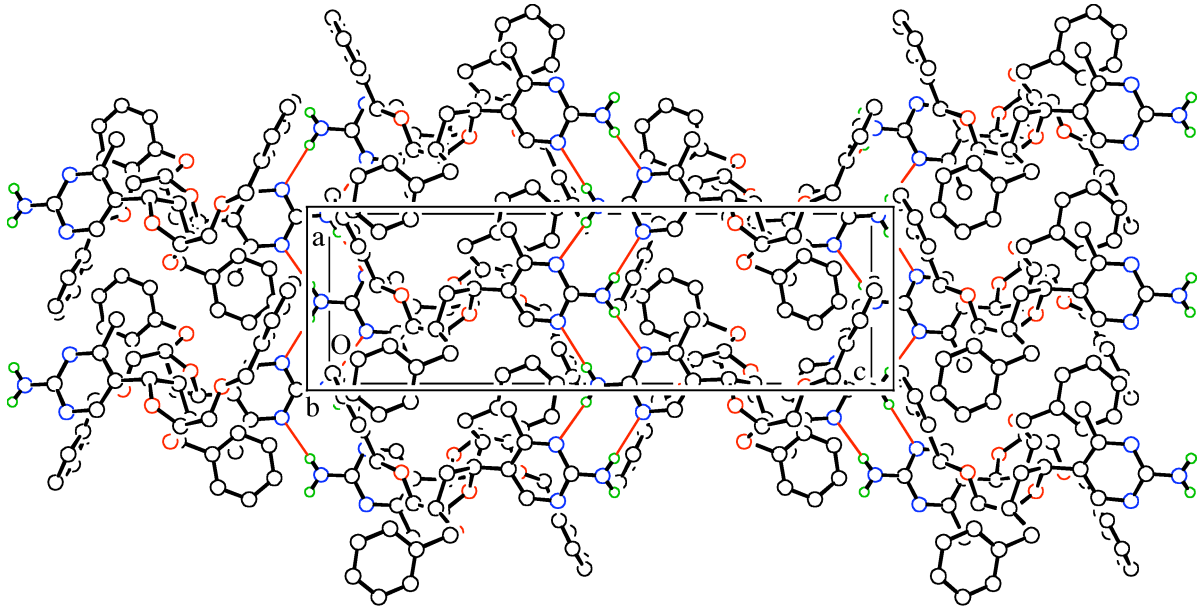


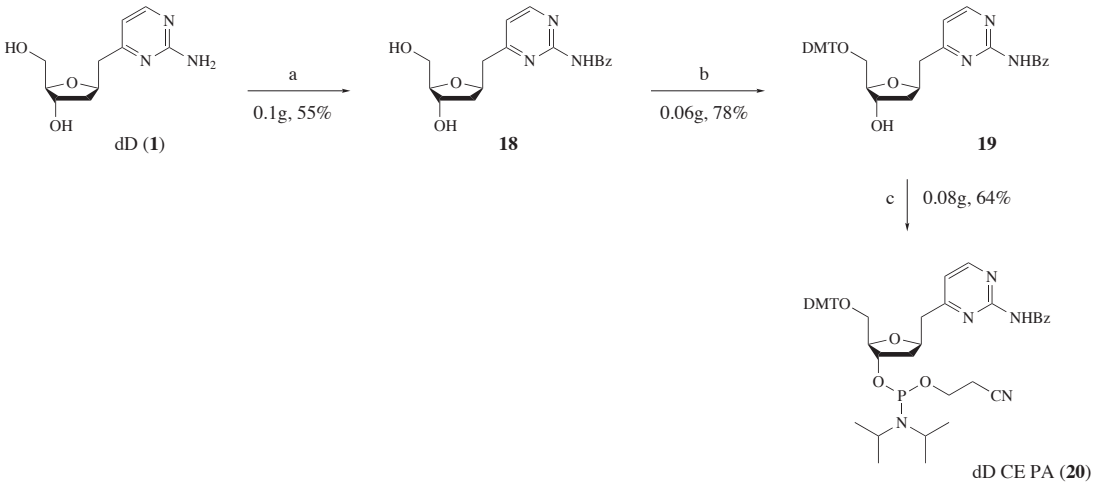
Figure 4.11: Packing of **13**α.

In the solid state, **13**α is found as a single conformer in which the sugar occurs in a C$^{2'}$-*endo* and the glycosidic bond in *anti* conformation. The torsion angle χ (O5–C1–C6–C7) is 153.9 ° (antiperiplanar). The three-dimensional structure is stabilized by hydrogen bonds; each aminopyrimidine ring forms four hydrogen bonds to two other nucleobases (see *Figure 4.11*).

From the NMR experiments of nucleoside **12**β and the observed coupling constants on can deduce rudimental information about its structure. The sugar pucker exists presumably in the C$^{2'}$-*endo* conformation, because a high ^3J coupling constant of 10.2 Hz was found between H$_\beta$C(2') and H$_\alpha$C(1'). This indicates roughly a H$_\beta$–C(2')–C(1')–H$_\alpha$ torsion angle of about 150°, which is only possible in the C$^{2'}$-*endo* conformation.[104] Between H$_\alpha$C(2') and H$_\alpha$C(1') a ^3J coupling constant of 5.2 Hz was found. The corresponding values observed for 2'-deoxyadenosine were 8.0 Hz (H$_\beta$C(2') – H$_\alpha$C(1')) and 6.0 Hz (H$_\alpha$C(2') – H$_\alpha$C(1')).[186] Additionally, one can assume that the torsion of the glycosidic bond is in the *anti* conformation, because no *NOE* correlations were found between the aromatic protons and HC(3') or H$_2$C(5').

Concerning the structural data, we can assume our nucleoside dD to be quite similar to its analog 2'-deoxyadenosine. Even though nucleoside dD is more flexible than dA, it is present in the *anti* conformation and should fulfill the fundamental structural requirements to be a possible analog of dA.

After successful separation and identification of our desired homo-*C* nucleoside **12**β, we carried on the synthesis by deprotection of the hydroxy groups giving dD (**1**). Thereafter, the fully unprotected nucleoside needed to be transformed to its cyanoethyl phosphoramidite (dD CE PA, **20**), which has been done by well-established chemistry in satisfying yields as seen in *Scheme 4.13*.[187] After selective monobenzoyl protection of the amino function, the 5'-hydroxy group was coupled with dimethoxytrityl (DMT), following a procedure described by *Eschenmoser* et al.[48] In a final step, the 3'-hydroxy group was converted to a cyanoethyl-*N*,*N*-diisopropyl phosphoramidite, namely dD CE PA (**20**).

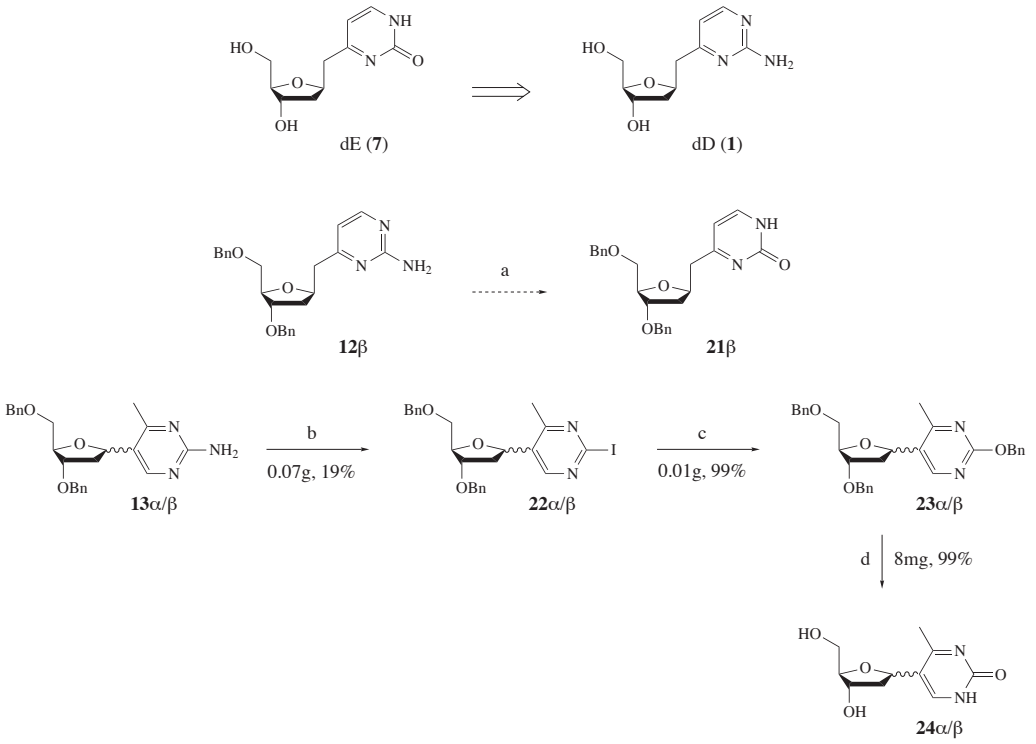


(a) i) TMSCl, Pyr, 0°C, 12h; ii) BzCl, rt, 12h. (b) DMTCl, $(Bu_4N)ClO_4$, Pyr, rt, 16h. (c) $PCl(N(i\text{-}Pr)_2)(OEtCN)$, $NEt(i\text{-}Pr)_2$, rt, 2h.

Scheme 4.13: Synthesis of dD CE PA (**20**).

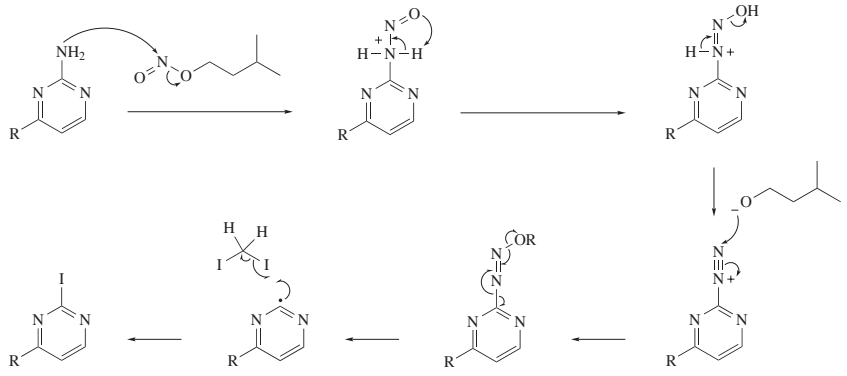
4.5 Attempted synthesis of dE

A second target molecule of this project was the G-analogue homo-*C* nucleoside dE (**7**). Following the original idea, which was outlined in *Scheme 4.5*, **7** should be synthesized from homo-*C* nucleoside dD (**1**). Because simple hydrolysis of the 2-aminopyrimidine under basic conditions was not successful, we decided to try another synthetic route. Due to the intricate separation of homo-*C* nucleoside **12**β, we tested this methodology with the undesired product *C*–nucleoside **13**α/β. The latter was first converted to its 2-iodopyrimidine analogue **22**, following a procedure by *Loren*[188] as seen in *Scheme 4.14*. A nucleophilic aromatic substitution with NaOBn led to compound **23**,[189] which was then fully deprotected in one step, giving the pyrimidone *C* nucleoside **24**. This trial showed that pyrimidones could be accessed by this methodology, but the first step, a radical iodization, worked only in poor yields. A proposed mechanism for this reaction is shown in *Scheme 4.15*.



(a) 10M NaOH, H_2O, 90°C, 2d. (b) $(C_5H_{11})ONO$, CH_2I_2, CuI, THF, 60°C, 1.5h. (c) NaOH, BnOH, MeCN, 85°C, 3.5h. (d) BBr_3, DCM, 0°C, 3min.

Scheme 4.14: Attempted synthesis of dE.



Scheme 4.15: Proposed mechanism.

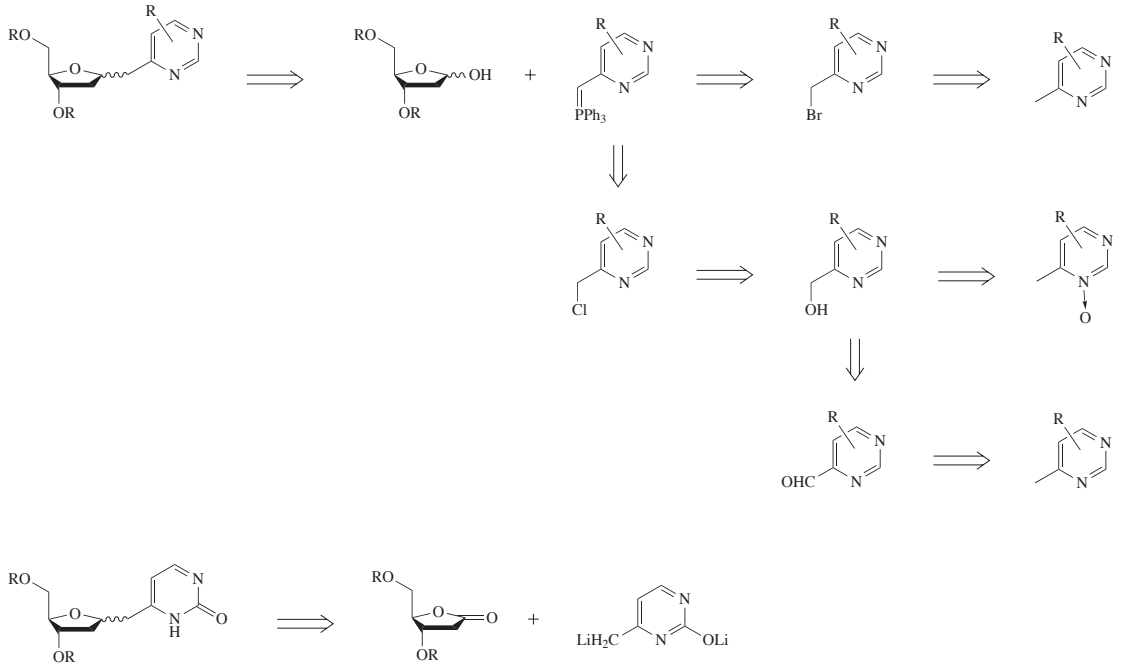
Various reaction conditions have been tested to improve the iodination of the 2-aminopyrimidine, but unfortunately yields up to only 20% were achieved (*Table 4.2*).[190–194] These results convinced us to work out a better strategy that allows us to access homo-*C* nucleoside dE easier and in larger scale.

Table 4.2: Tested conditions for iodination of 2-aminopyrimidine **13**.

Reaction conditions	Yield [%]
$(C_5H_{11})ONO$, Benzene, I_2, 80°C, 36h	–
$(C_5H_{11})ONO$, CH_2I_2, 85°C, 3.5h	11
$(C_5H_{11})ONO$, CH_2I_2, CuI, I_2, THF, 66°C, 0.5h	18
$(C_5H_{11})ONO$, CH_2I_2, CuI, THF, 60°C, 1.5h	19
CH_2I_2, CuI, THF, 66°C, 1.5h	14–20

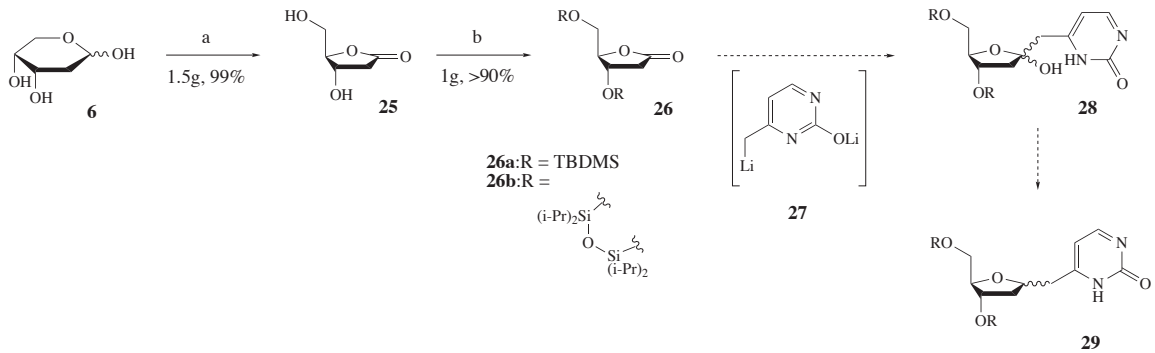
Scheme 4.16 outlines two possible alternative strategies, which have been tried to synthesize homo-*C* nucleoside dE. The upper one targets toward a *Wittig*-type reaction to build up the homo-*C* glycosidic bond, a method which has been developed and used by *Townsend*.[168] The key intermediate of this method would be a *Wittig* reagent of the suitable substituted 4-methylpyrimidine.[195] Unfortunately, we have not been able to synthesize such a compound, although we tried several different pathways and derivatives. The classical approach *via* a 4-(bromomethyl)pyrimidine failed with various substitution patterns as well as under a series of conditions because bromination of the corresponding 4-methylpyrimidine always yielded 5-bromo-4-methylpyrimidine. Also, a second approach, starting with the *N*-oxide of a 4-methylpyrimidine,[196–199] which could be used in a *Böckelheide* reaction to further functionalize the methyl group, failed in our hands.

A more promising strategy is shown at the bottom of *Scheme 4.16*. The fundamental idea is to build up the glycosidic bond by a nucleophilic attack of a double lithiated 4-methylpyrimidone to the lactone of deoxyribose.[200] As seen in *Scheme 4.17*, lactone **26** was synthesized with two different silyl protecting groups, TBDMS (**26a**) and bridging tetraisopropyldisiloxandiyl(**26b**),[162,201,202] according to the literature and was, thereafter, coupled with double-lithiated 2-hydroxy-4-methyl-pyrimidine **27**.[203] Instead



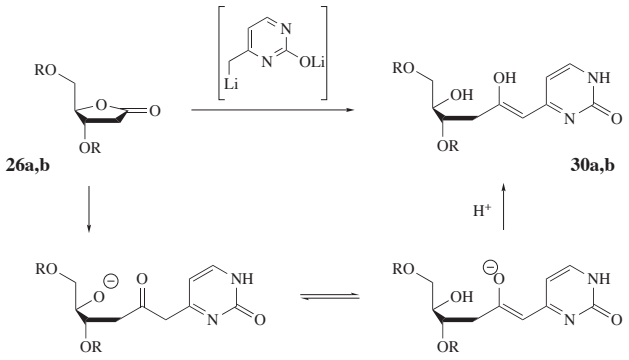
Scheme 4.16: Alternative strategies.

of desired product **28**, which should be converted to homo-*C* nucleoside **29**, compound **30** was formed in reactions under various conditions and with both silyl protecting groups. **30** can be explained by a sugar ring opening during nucleophilic attack, resulting in a enol that is stabilized by conjugation with the pyrimidin-2(*1H*)-one (see *Scheme 4.18*).



(a) Br_2, H_2O, 0°C, 20h. (b) RCl, DMF, rt, 20h.

Scheme 4.17: Attempted synthesis of dE.



Scheme 4.18: Sugar ring opening.

These results were disappointing because such ring opening in similar reactions is, to the best of our knowledge, not described in the literature, but the methodology seems to be promising and would allow to access homo-*C* nucleoside dE in only four steps. Ring opening could possibly be avoided by insertion of a good leaving group at C(1), such as the C(1)–Cl compound.[204]

4.6 Base Pairing Studies

In order to investigate the potential of our artificial nucleotide **dD** to mimic the natural nucleobase **dA**, we incorporated it into oligonucleotides and to perform base-pairing studies. We decided to use a system that was developed and investigated, in another contex, by Breslauer. He used a non-self-complementary 13^{mer} oligonucleotide (see *Figure 4.12*) to study the thermodynamic contribution that DNA single-stranded order makes to DNA duplex formation.[205–207]

5'– C G C A U G A G U A C G C –3'
⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮
3'– G C G U A C U C A U G C G –5'

Figure 4.12: Investigated complementary oligonucleotides.

5'– C G C A U G **X** G U A C G C –3'
⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮
3'– G C G U A C **Y** C A U G C G –5'

Figure 4.13: Positions to modify in the complementary strands.

4.6.1 Synthetic Constructs – 13-mers

We planned to synthesize the two complementary strands and to modify the highlighted middle positions **X** and **Y**. dA, dG, dC, dU, and dD should be incorporated at these positions in both strands. Accordingly, we had to synthesize ten different strands, five of each kind. To simplify matters, we will call the two different kinds of strands "**A**" and "**B**" and number them according to *Table 4.3*.

Our intention was to combine the complementary strands in all 25 possible combinations and perform melting experiments with these double-stranded oligonucleotides in order to obtain a complete set of thermodynamic data and to determine the base-pairing abilities of homo-*C* nucleoside dD compared to all other bases.

Table 4.3: Abbreviations for the synthesized strands.

Kind of strand	Oligonucleotide	Abbreviation
Strand A	5'–CGCAUG**A**GUACGC–3'	A1
	5'–CGCAUG**D**GUACGC–3'	A2
	5'–CGCAUG**G**GUACGC–3'	A3
	5'–CGCAUG**U**GUACGC–3'	A4
	5'–CGCAUG**C**GUACGC–3'	A5
Strand B	5'–GCGUAC**A**CAUGCG–3'	B1
	5'–GCGUAC**D**CAUGCG–3'	B2
	5'–GCGUAC**G**CAUGCG–3'	B3
	5'–GCGUAC**U**CAUGCG–3'	B4
	5'–GCGUAC**C**CAUGCG–3'	B5

Oligonucleotides were synthesized on an automated DNA synthesizer using the *Phosphoramidite* approach and purified as explained in details in *Sections 3.5* and *6.20*. *Figure 4.14* shows the purification of crude oligonucleotide B2 by HPLC. Due to incomplete couplings, oligonucleotides of reduced length are present in the raw product as seen in the upper chromatogram.[2] The lower chromatogram exhibits the same oligomer after separation on an analytical *DEAE*-ion exchange HPLC column.

[2]Chromatogram shifted up by 1 mAU for better clarity

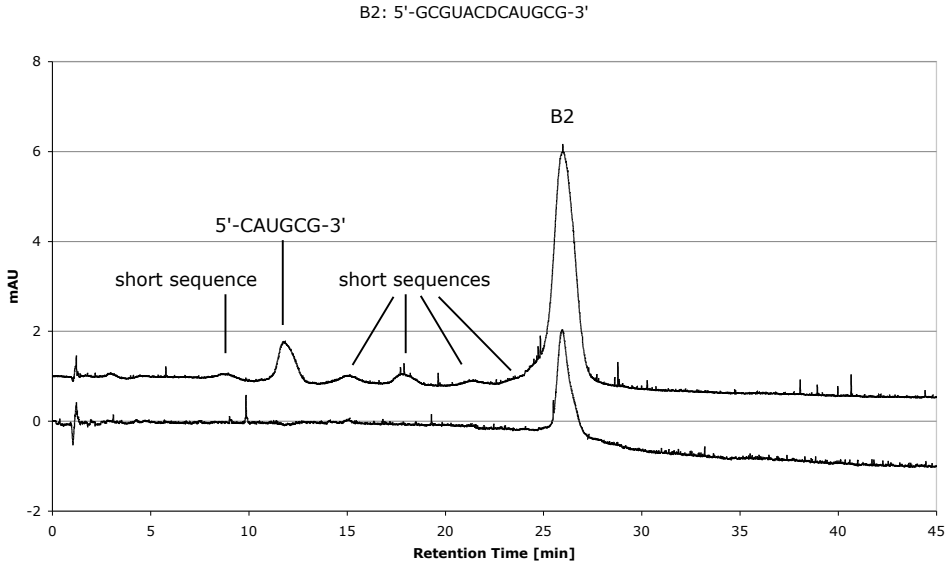
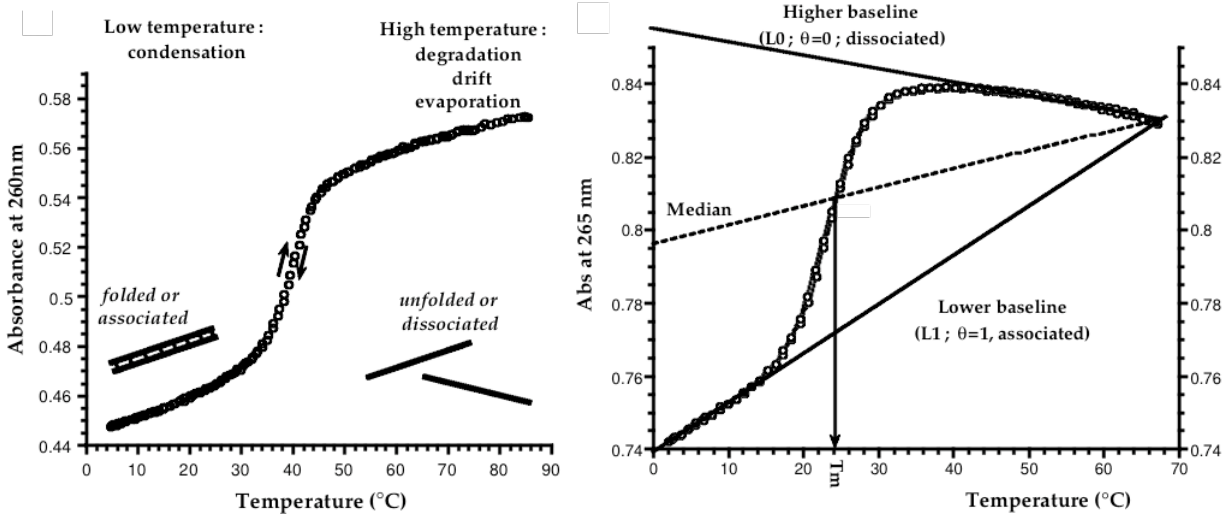


Figure 4.14: Purification of crude oligonucleotide B2 by HPLC.

4.6.2 Melting Studies

After purification, the oligonucleotides were used to measure melting curves of double-stranded DNA by temperature-dependent UV-spectroscopy. Complementary sequences were combined in a 1:1 ratio in buffer solution and slowly heated up recording the absorbance at 260 nm as a function of temperature. The principle of such an experiment was developed by *Thomas*[208] more than 50 years ago and is based on hypochromicity – an effect which is observed in oligonucleotides as well as in other (bio-)polymers.[209] Double-stranded DNA with a well defined three dimensional structure absorbs less light than single-stranded DNA. Accordingly, melting of DNA, the change from an associated duplex at low temperature to unfolded single strands at high temperature, can be followed by an increase of UV absorbance (see *Figure 4.15a*). Duplex denaturation of DNA leads to a hyperchromism of 15 % – 20 %.[210] In principle, such a thermal denaturation curve could also be obtained with fluorescence emission spectroscopy,[211] NMR measurements,[212] circular dichroism,[213] or *Raman* spectroscopy.[214]



(a) Exemplary melting curve *(b)* Determination of T_m

Figure 4.15: Melting curves and determination of T_m.[210]

From temperature dependent melting curves one can determine the melting temperature (T_m) of a given double-stranded DNA (see *Figure 4.15b*). Its T_m corresponds to the temperature at which half of the sample is folded (double-stranded), and half is unfolded (single-stranded). Several parameters influence the strength of duplex binding and therefore the melting temperature T_m.[210] First, the length of the oligonucleotide affects the T_m. Second, the composition of the complementary oligonucleotides makes a difference in T_m, which increases linear with the molar content of G-C base pairs. Third, in some cases T_m is dependent on the oligonucleotide concentration. In the case of an intramolecular complex (e.g. *i*-Motif), the T_m should be concentration independent, whereas in the case of bimolecular (autocomplementary or two different complementary oligonucleotides) or multimolecular complexes T_m should be dependent on the concentration. Additionally, other parameters like, ionic strength, the nature of the salt, or the pH influence the melting temperature. An increase in ionic strength leads to an increase of the T_m, because the cation concentration will partially neutralize the net charge of the nucleic acid, which is a negatively charged polyelectrolyte (see *section 3.3.2.5*).[215] Moreover, pH plays an important role in the lability of DNA in

consideration of the protonation of nucleobases.

From a measured melting curve one can now deduce the thermodynamic parameters of the melting process and therefore quantify the strength of duplex binding. Several different methods to calculate the thermodynamic data are common and will be explained in details later in this section.

A melting experiment needs to be set up carefully. For instance, buffer solutions and samples should be degassed in order to avoid air bubbles, and parameters of the measurement should be examined precisely. The heating temperature gradient is a crucial point. A temperature slope that is too steep would lead to an inaccurate result, because the thermal equilibrium between double- and single-stranded DNA is not reached. *Figure 4.16* shows a typical melting curve, recorded with a temperature gradient of 12 °C per hour. A *reverse temperature* (RT) scan does not differ significantly, but a 24 °C/h scan is shifted considerably. This observation indicates that 12 °C is an adequate gradient, which allows the sample to reach its thermal equilibrium but, on the contrary, heating the sample by 24 °C/h is too fast and results in a wrong melting temperature (T_m).[210]

Melting curves were measured for four different oligonucleotide concentrations in each case as seen in *Figure 4.17*. Because T_m is concentration-dependent in the case of a bimolecular association of two independent non-self-complementary strands, at least three measurements are necessary to calculate the thermodynamic parameters ΔG°, ΔH° and ΔS°. The interpretation of the achieved data and calculations of the thermodynamic parameters were performed in two different ways in order to verify the results. First, melting curves were analyzed graphically in an absorbance *vs.* temperaure plot to determine the melting temperatures (see *Figure 4.15a*). Furthermore, a *fraction folded* plot was converted from the absorbance *vs.* temperature plot to approve this values. A *fraction folded* plot is a representation of *fraction folded* (θ) *vs.* temperature. The lower baseline in the Absorbance *vs.* Temp. plot corresponds to the associated, double stranded duplex ($\theta = 1$, $T \ll T_m$),

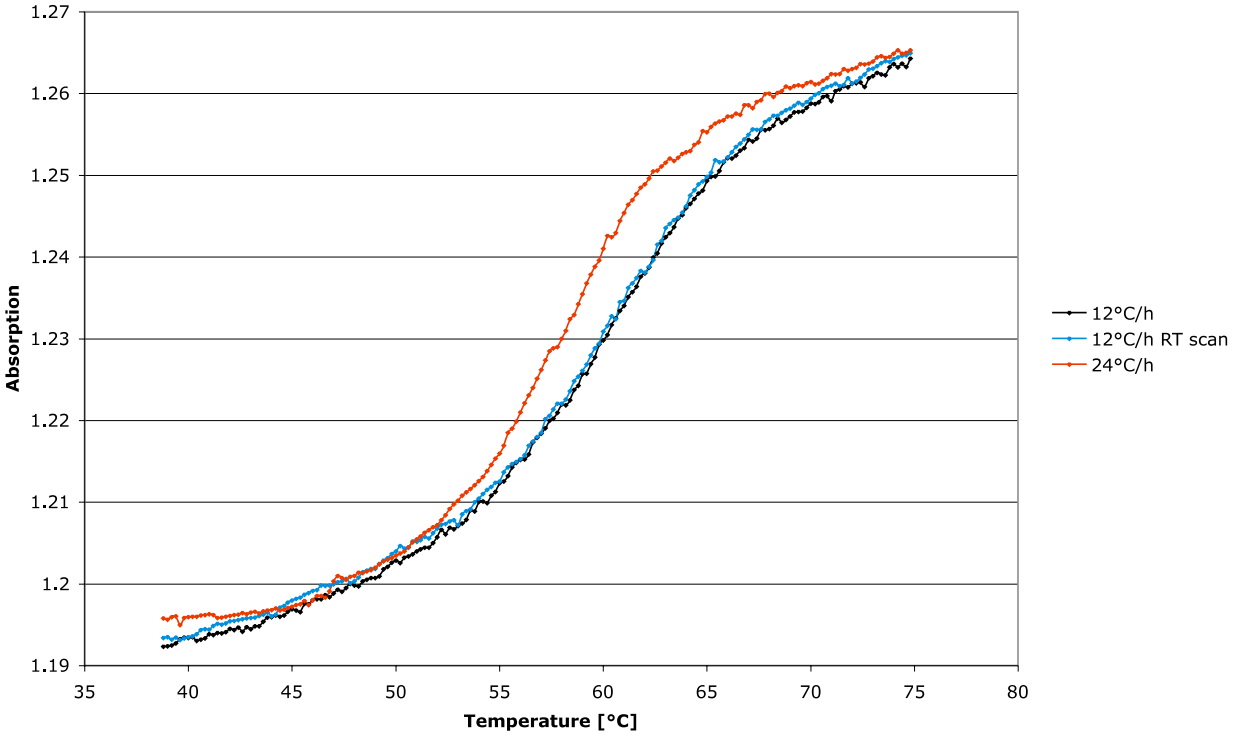


Figure 4.16: Testing for an adequate temperature slope.

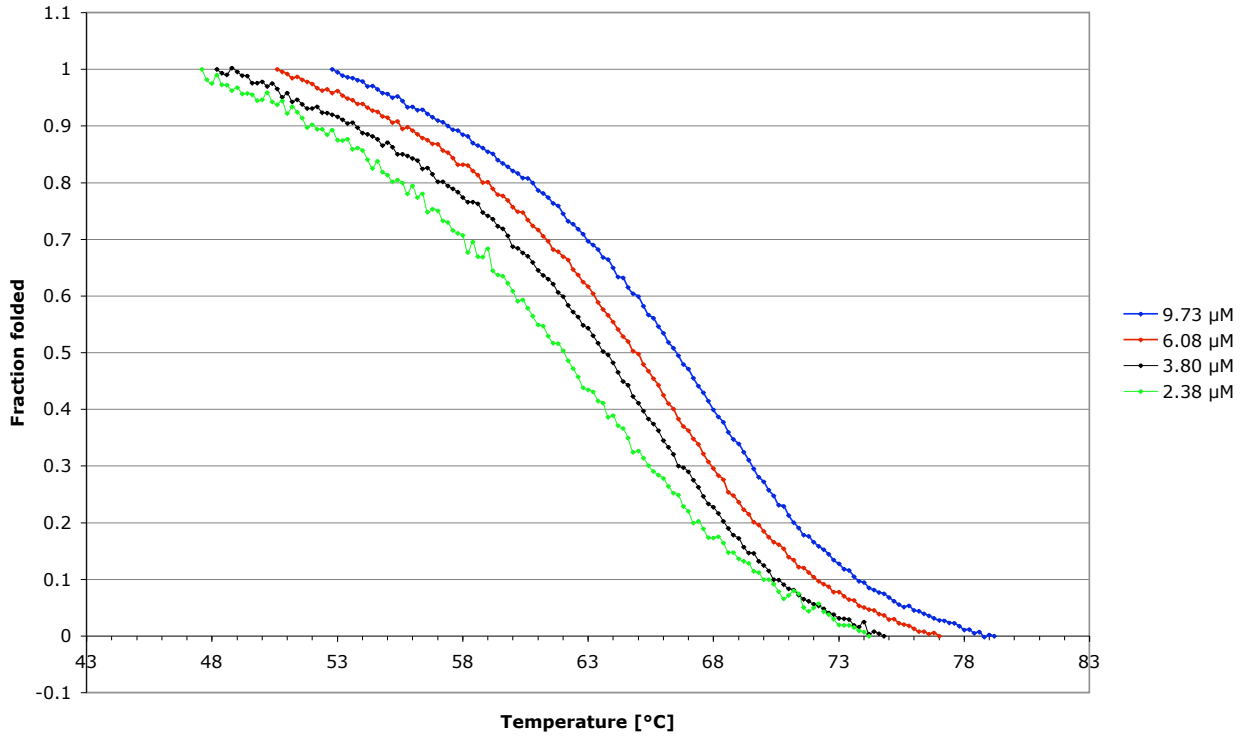


Figure 4.17: Example of a fraction folded plot with four different oligonucleotide concentrations.

whereas the upper baseline corresponds to the dissociated, single stranded oligonucleotides ($\theta = 0$, $T \gg T_m$). Accordingly, the melting temperature $T = T_m$ corresponds to $\theta = 0.5$. A *fraction folded* plot is derived by the *Formula 4.1*, whereat $L0_T$ and $L1_T$ correspond to the baseline absorbance values of the unfolded and folded species, and A_T is the absorbance at the given temperature.

$$\theta_T = \frac{L0_T - A_T}{L0_T - L1_T} \tag{4.1}$$

With the obtained T_m values, one can now calculate the *Van't Hoff* pairing standard ethalpies ΔH° and standard entropies ΔS° according to described methods.[216,217] By plotting $^1/T_m$ as a function of the initial oligonucleotide concentration (C_{Total}), one may then write:

$$\frac{1}{T_m} = \frac{\Delta S^\circ}{\Delta H^\circ} + \frac{R * ln(C_{Total}/2)}{\Delta H^\circ} \tag{4.2}$$

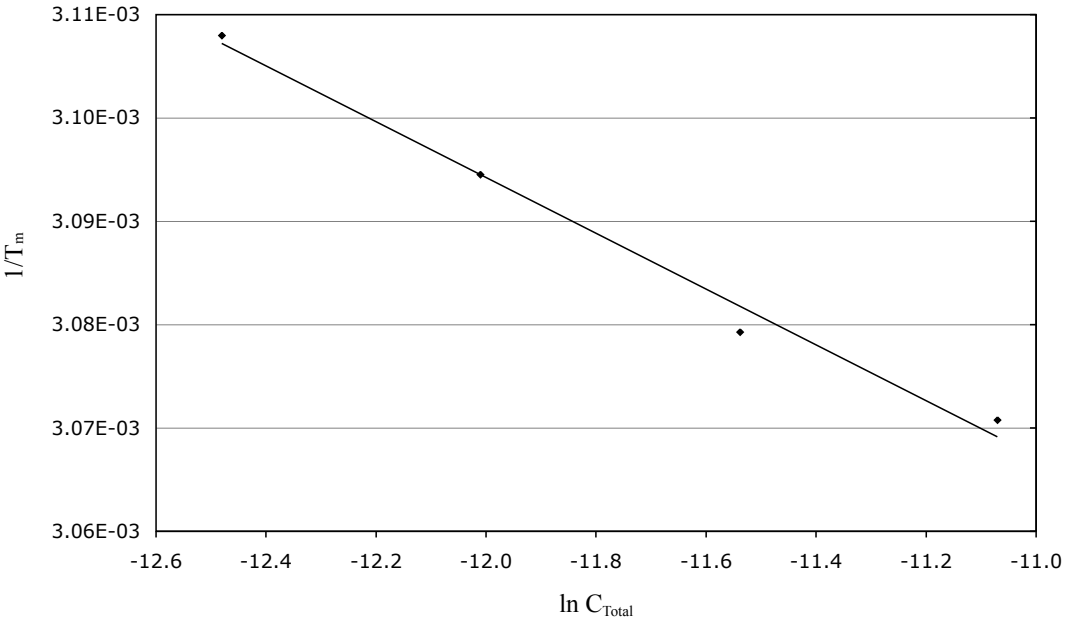


Figure 4.18: Plot of $^1/T_m$ vs. ln C_{Total}.

Out of this correlation, one can deduce ΔH° and ΔS° as explained in *Figure 4.18* and *Formulas 4.4*, *4.5*. Finally, one can calculate the *Gibbs* free enthalpy, which, by definition, may be written as:

$$\Delta G^\circ = \Delta H^\circ - T * \Delta S^\circ \tag{4.3}$$

$$slope = R/\Delta H^\circ \tag{4.4}$$

$$intercept = \frac{\Delta S^\circ - Rln4}{\Delta H^\circ} \tag{4.5}$$

These parameters can also be determined from a second plot: the *Van't Hoff* representation, a correlation of ln K_a vs. $^1/_T$ (see *Figure 4.19* and *Formulas 4.7*, *4.8*). K_a, the affinity constant, can be calculated from the fraction folded θ_T according to *Formula 4.6*, where C_0 is the single-strand concentration and one should generally restrict the analysis to the temperature range for which $0.15 < \theta < 0.85$.[210]

$$K_a = \frac{\theta_T}{C_0 * (1 - \theta_T)^2} \tag{4.6}$$

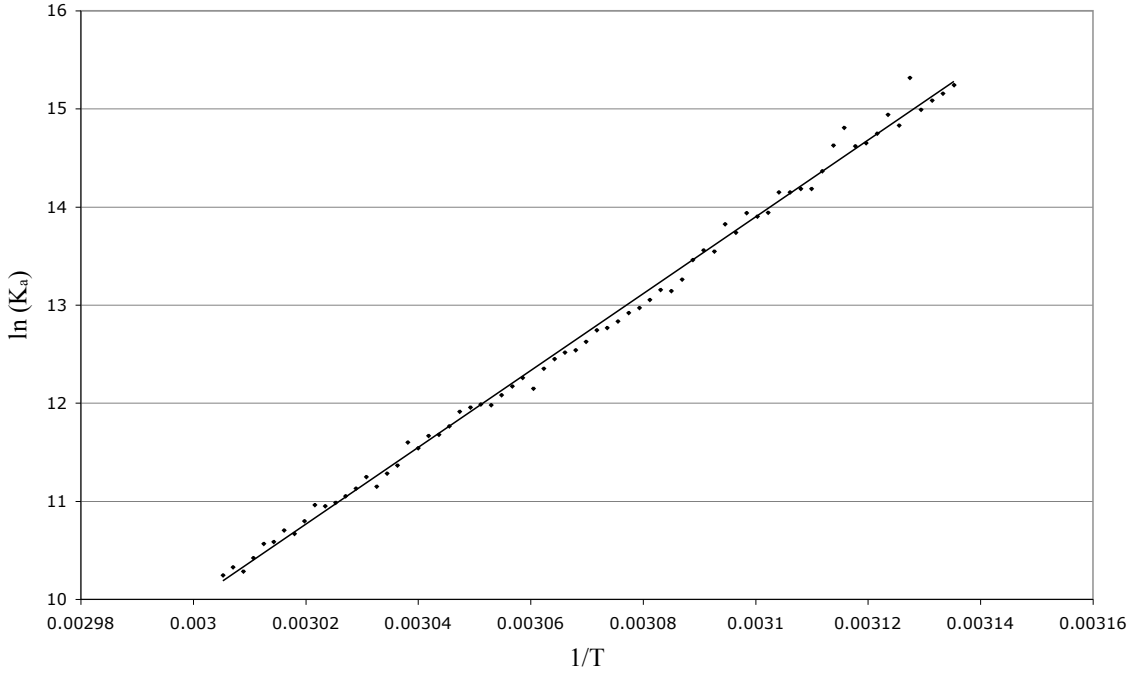


Figure 4.19: Van't Hoff plot: ln K_a vs. $^1/_T$.

$$slope = -\frac{\Delta H^\circ}{R} \tag{4.7}$$

$$intercept = \frac{\Delta S^\circ}{R} \tag{4.8}$$

Both discussed representations to determine the thermodynamic parameters require ΔH and ΔS to be temperature-independent, an assumption that is not always valid but as described by *Chaires* very difficult to analyze.[218]

The resulting thermodynamic parameters should be equal for both calculations and can be compared. Problems can occur if the baselines are not set exactly, for instance. Such a misinterpretation has a decisive effect on the outcome of the $^1/_{T_m}$ vs. ln C_{Total} calculation, but can be perceived because it induces a significant curvature of the *Van't Hoff* plot.

On the following page, all accumulated thermodynamic data are summarized. I will integrate this comprehensive set of data and draw conclusions that are valid for this special case of double-stranded DNA as well as characterize the new artificial nucleotide dD in general.

Table 4.4: Summary of ΔG° values [$^{kcal}/_{mol}$].

		Strand A				
		A	D	G	U	C
Strand B	A	-14.9	-12.7	-14.2	-18.1	-13.6
	D	-13.9	-11.8	-14.2	-15.2	-12.8
	G	-12.6	-11.4	-13.8	-15.5	-17.5
	U	-18.5	-12.6	-16.0	-14.6	-14.0
	C	-13.0	-11.4	-18.2	-13.5	-13.2

Table 4.5: Summary of ΔH° values [$^{kcal}/_{mol}$].

		Strand A				
		A	D	G	U	C
Strand B	A	-76.8	-57.0	-64.9	-88.6	-70.8
	D	-76.8	-56.1	-77.1	-80.9	-70.3
	G	-57.5	-54.2	-67.6	-84.1	-79.4
	U	-94.7	-54.3	-87.9	-77.9	-78.5
	C	-62.7	-48.3	-86.3	-70.5	-76.8

Table 4.6: Summary of ΔS° values [$^{cal}/_{mol*K}$].

		Strand A				
		A	D	G	U	C
Strand B	A	-208	-148	-170	-237	-192
	D	-211	-149	-211	-220	-193
	G	-151	-144	-181	-230	-208
	U	-256	-140	-241	-212	-216
	C	-167	-124	-229	-191	-213

Table 4.7: Comparison of dA and dD by ΔΔG° values [kcal/mol].

		Strand A		
		A	D	ΔΔG°
Strand B	A	-14.9	-12.7	2.2
	D	-13.9	-11.8	2.1
	G	-12.6	-11.4	1.2
	U	-18.5	-12.6	5.9
	C	-13.0	-11.4	1.6

		Strand B		
		A	D	ΔΔG°
Strand A	A	-14.9	-13.9	1.0
	D	-12.7	-11.8	0.9
	G	-14.2	-14.2	0.0
	U	-18.1	-15.2	2.9
	C	-13.6	-12.8	0.8

Table 4.8: Comparison of dA and dD by ΔΔH° values [kcal/mol].

		Strand A		
		A	D	ΔΔH°
Strand B	A	-76.8	-57.0	19.8
	D	-76.8	-56.1	20.7
	G	-57.5	-54.2	3.3
	U	-94.7	-54.3	40.4
	C	-62.7	-48.3	14.4

		Strand B		
		A	D	ΔΔH°
Strand A	A	-76.8	-76.8	0.0
	D	-57.0	-56.1	0.9
	G	-64.9	-77.1	-12.2
	U	-88.6	-80.9	7.7
	C	-70.8	-70.3	0.5

Table 4.9: Comparison of dA and dD by ΔΔS° values [cal/mol∗K].

		Strand A		
		A	D	ΔΔS°
Strand B	A	-208	-148	60
	D	-211	-149	62
	G	-151	-144	7
	U	-256	-140	116
	C	-167	-124	43

		Strand B		
		A	D	ΔΔS°
Strand A	A	-208	-211	-3
	D	-148	-149	-1
	G	-170	-211	-41
	U	-237	-220	17
	C	-192	-193	-1

4.6.3 Trends in ΔG° of pairing

As expected, we have found the highest binding energies in the cases of the natural *Watson-Crick* base pairs, A:U and G:C, where we observed ΔG°, ΔH°, and ΔS° values of -17.5 – -18.5 $^{kcal}/_{mol}$, -79.4 – -94.7 $^{kcal}/_{mol}$ and -208 – -256 $^{cal}/_{mol*K}$, respectively. These values are comparable to the results measured by *Breslauer* with the similar 13-mer containing A:T instead of A:U base pairs.[219] He observed $\Delta G^\circ = 20.0$ $^{kcal}/_{mol}$, $\Delta H^\circ = 117.0$ $^{kcal}/_{mol}$, and $\Delta S^\circ = 325.4$ $^{cal}/_{mol*K}$ in 10 mM sodium phosphate buffer solution[3].

5'– C G C A T G A G T A C G C –3'
⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮ ⋮
3'– G C G T A C T C A T G C G –5'

Figure 4.20: 13-mer duplex investigated by Breslauer.

Several properties of the particular single strands and of the double stranded DNA effect these results and are responsible for the occasionally observed differences. Simplified, ΔG° is in general the sum of the binding between the two strands (ΔH°, strongly negative: highly favorable) and the stacking of the strands (ΔS°, very negative: very unfavorable). As one can imagine, these two parameters are not independent from each other, but influence one another. The event of duplex formation from two complementary "random coil" single strands can be abstractly divided into two parts. First, every single strand forms a pre-ordered, stacked structure. Then, the pre-stacked single strands bind to each other and form the duplex. *Breslauer* has observed that the two complementary single strands already posses $> 40\%$ of the total enthalpy.[219] Certainly, duplex formation cannot be considered as an explicit three-state process, but for better clarity, this is a descriptive model. Obviously, an optimal prestacking facilitates an effective binding, whereas a disrupted prestacking disturbs the binding affinity. This interrelation makes

[3] 10 mM sodium phosphate, 1 mM EDTA, 1.0 M NaCl, adjusted to pH 7.

it difficult to draw accurate conclusions about the binding properties of a single base pair because its influence is spread out over the whole duplex structure. However, general trends can be identified and will be outlined in the following discussion.

By comparing the thermodynamic data of dA *versus* dD against all other natural nucleobases, a general trend was observed. Binding, as well as prestacking, seems to be weaker in oligonucleotides containing dD compared to those containing dA. This observation is not surprising due to the fact that the glycosidic methylene bridge in homo-*C* nucleosides introduces an additional degree of freedom into the sugar–nucleobase junction (see *Figure 4.21*).

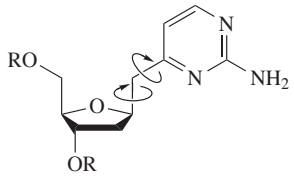


Figure 4.21: Additional degree of freedom in nucleoside dD.

The extra flexibility presumably disrupts the base stacking and weakens the ability of binding to its counterpart. Furthermore, the aminopyrimidine ring of nucleoside dD is probably less polarizable than its purine analogue dA and, therefore, less compatible for binding and stacking. Pyrimidines generally stack less strongly than purines, presumably because purines have larger surface area and greater polarizability.[115]

We observed a slight selectivity of dD to bind to its target complementary base dU. Unfortunately, binding to the nucleobase dA was also found to be in the same range of strength as, in one case, to nucleoside dG (see *Figure 4.22*). If nucleoside dD in strand A (A2) is merged with all complementary stands B, the combinations D:A (A2B1) and D:U (A2B4) expose the highest affinities of about $\Delta G^\circ \approx 12.6$ $^{kcal}/_{mol}$. Although the values are lower than in the dA series (A1, $\Delta G^\circ \approx 15$ $^{kcal}/_{mol}$), a clear similarity can be noticed.

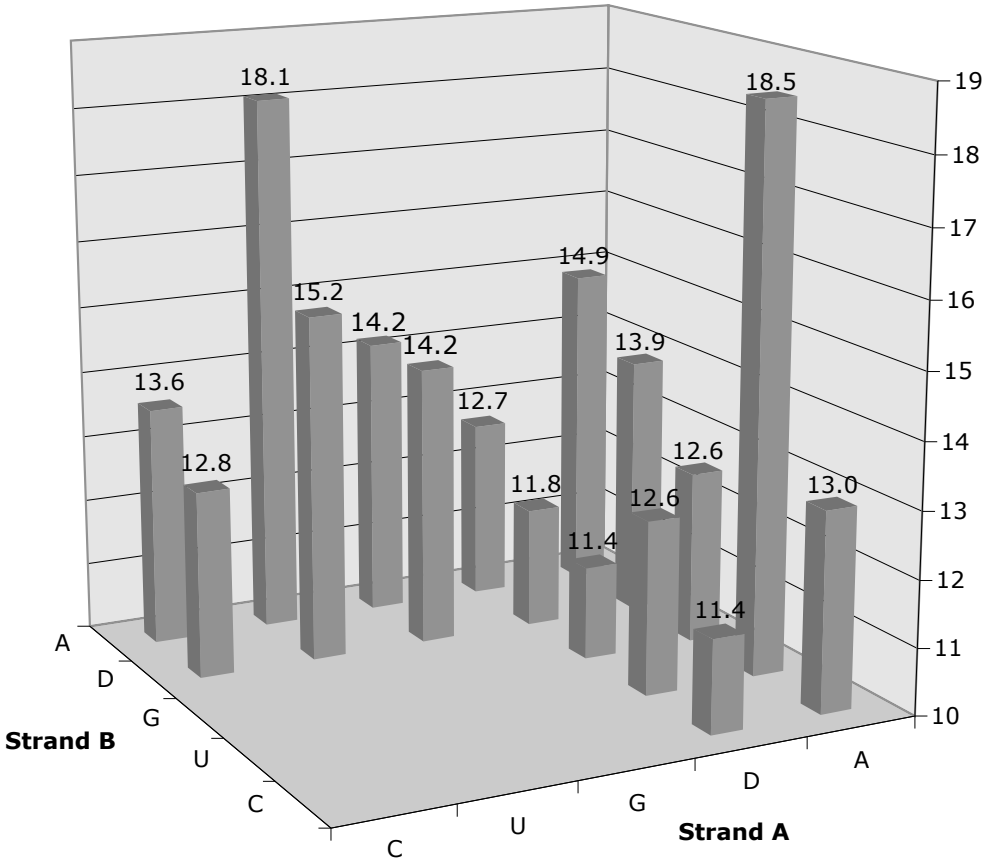


Figure 4.22: -ΔG° values [$^{kcal}/_{mol}$] of the dA and dD series.

This decrease in binding affinity to their complementary bases may be explained by the different stacking abilities of dA and dD. In strand A, the middle position is flanked by two guanosines. As known in the literature, dA and dG have the highest stacking free energies of the four natural nucleosides and can therefore form stable stacked structures when they occur in a row as in strand A1 (see *section 3.3.2.5*).[115] On the other hand, the stacking is disrupted if nucleoside dD occurs in the middle position, as in strand A2, and the binding affinity decreases. ΔG° values of strand A1 against all other complementary strands were in the range of -12.6 – -18.5 $^{kcal}/_{mol}$, whereas the valus of strand A2 were in the range of -11.4 – -12.7 $^{kcal}/_{mol}$. In general, all melting studies with A1 and A3, oligonucleotides with three purines in a row located in the middle, showed highly negative ΔS° values and support this argument.

The fact that dD forms somewhat stable duplexes with dU, as well as with dA, allows the assumption that hydrogen bond motifs between these nucleobases as outlined in *Figure 4.23* are present. The structural changes needed to arrange the proposed bonding between dD and dA, for instance, possibly does not dramatically disturb the stacking.

If we compare the strands B1 and B2, which accommodate dA and dD in the middle positions, respectively, the trend, that the highest ΔG° values are found if dD pairs with dU and dA, is also observed. In contrast to the A2 series, quite high ΔS° values are found in the B2 series, which indicate a quite stable stacking of dD with its flanking cytidine nucleobases in strand B. C is known to induce strong dipoles in its nearest neighborhood and may fix the flexible homo-*C* nucleotide dD in a stable, stacked conformation.

In general, all observed thermodynamic data demonstrate that dD is much more similar to dA in strand B than in strand A. $\Delta\Delta G^\circ$, $\Delta\Delta H^\circ$, and $\Delta\Delta S^\circ$ are much smaller when we compare dD and dA in strand B than the corresponding data of strand A. The flanking cytidines seem to be able to influence nucleotide dD in a manner that allows it to stack and pair quite analogous to dA. Such an amplitude of a nearest neighborhood effect was absolutely not expected and the question of how other flanking nucleotides would influence the binding affinity of homo-*C* nucleotide dD arises. Unfortunately, we will not be able to answer this problem in this thesis.

A strong binding was also found in this series in the case where dD pairs with dG (B2A3, $\Delta G^\circ = 14.2\ ^{\text{kcal}}/_{\text{mol}}$). On one hand, the already mentioned stacking effect of three purines in a row in strand A3 could be responsible for the high ΔS° value. But, on the other hand, a possible hydrogen bond motif, as seen in *Figure 4.23*, can also be proposed based on the observed high ΔH° value.

Another exceptional effect we have observed, was a strong affinity of strands were dG is in line with dU (A3B4 and A4B3, $\Delta G^\circ \approx 16$ $^{\text{kcal}}/_{\text{mol}}$). Since a hydrogen bond motif can be excluded, stacking forces presumably are responsible for this affinity.

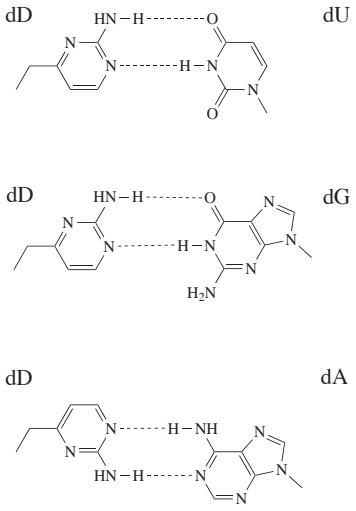


Figure 4.23: Possible binding patterns of dD with dU, dG and dA.

4.7 Outlook

From the insights we gained in the field of homo-*C* nucleotides and the aminopyrimidine nucleotide dD in particular, countless new questions and very interesting scientific challenges have arisen. As discussed in the last chapter, homo-*C* nucleotide dD is partially able to mimic the natural nucleotide dA and, therefore, these results open a completely new area of scientific research. First of all, some open questions concerning the binding and stacking ability of dD need to be answered. The nearest neighborhood effect of dU, dA, and dD against dD must be investigated. In my opinion, the non-self-complementary oligonucleotides that we have investigated in this work are very suitable for such a disquisition. By varying the positions that flank the nucleotide in the middle, where a purine should be exchanged for a purine and a pyrimidine for a pyrimidine, the nearest neighborhood effect could be determined qualitatively. Toward the same aim, a 5'-dC-[dD-dC]$_n$-3' strand, and its complementary strand 3'-dG-[dU-dG]$_n$-5', should be synthesized to verify the result that nucleotide dC is able to stabilize dD in a stacked conformation.

Furthermore, homo-*C* nucleotide dE should be synthesized and investigated in the same oligonucleotide system. Presumably, it shows higher binding and stacking abilities than dD, due to its higher molecular dipole which could also lead to a higher nucleobase selectivity in binding.

At the next level, an all-pyrimidine system should be evaluated in order to test the hypothesis of *Siegel* and *Tor*,[1] which suggests an all-pyrimidine progenitor of present-day DNA.

If we look further ahead, the possible range of application for such homo-*C* nucleosides is unlimited. The effect of a dA to dD exchange could be investigated in every context where adenosine occurs and participates in biochemical processes. Possible targets could be ATP, NAD^+, $NADP^+$, FAD, or CoA, for instance. As observed for some EANs and *C*-nucleotides, homo-*C* nucleosides probably show antiviral or antibiotic activities (*see sections 3.4 and 3.4.3*).

Experimental Part

Chapter 5

General

All reagents were used as purchased from commercial suppliers unless otherwise stated. Solvents were purified and dried by conventional methods prior to use. Thin-layer chromatography (TLC): *Merck* TLC aluminium sheets, silica gel *60* F_{254}, 2mm. Column chromatography (CC): $Sigma - Adrich$ Silica gel Merck Type 9385, 230-400 mesh, 60 Å. High-performance liquid chromatography (HPLC): Analytical HPLC: *Shimadzu LC–10AT*, Spherisorb®-NH_2, 5 μm, 25cm×4.6mm. Preparative HPLC: *Shimadzu LC–8A*, Spherisorb®-NH_2, 5 μm, 25cm×20mm. IR: $Perkin - Elmer -$ 1600 FT-IR spectrometer; absorption values in cm^{-1} and intensity (*s*: *strong*, 0-30% transmission; *m*: *middle*, 30-60% transmission; *w*: *weak*, 60-100% transmission).^{1}H- and ^{13}C-NMR: *Bruker ARX–300*, *AV–400*, *DRX–500* or *AV–600* instruments. ^{13}C-signal multiplicity was deduced from DEPT 90 and DEPT 135 spectra (*Disortionless Enhancement by Polarization*). Peak assignment was performed by two-dimentional NMR experiment (NOESY, COSY, HSQC, HMBC). MS: *Finnigan MAT–95* instrument for CI and EI; *Finnigan TSQ–700* triple quadrupole spectrometer for ESI; m/z (rel. %); *Bruker* Autoflex I spectrometer for MALDI-TOF. EA: *Vario EL* instrument (*Elementar*). UV and UV-melting curves: *Jacso* J-715 spectrapolarimeter equipped with a *Julabo* HS 18 temperature control unit in buffered solution (1 M NaCl, 0.01 M TrisHCl buffer, pH = 7.0)

Chapter 6

Synthesis

6.1 Methyl-3,5-di-*O*-benzyl-2-deoxy-α/β-D-ribofuranosid (8)

According to the literature,[178,220–222] **8** was synthezised from 2-Deoxy-D-ribose in good yields. An anomeric ratio of α/β 1.2:1 was observed. Anomers could be separated by CC (SiO_2, Hex/Et_2O 1:1) and identified by NOESY NMR.[1]

6.2 Methyl-3,5-di-*O*-benzoyl-2-deoxy-α/β-D-ribofuranosid (14)

According to the literature,[178] Methyl-2-deoxy-α/β-D-ribofuranosid **5** was synthesized from 2-deoxy-D-ribose. 1.007 g **5** (7.508 mmol) was dissolved in 5 ml Pyr, cooled to 0 °C, and ,after adding BzCl (2.44 ml, 0.021 mol, 2.8 eq.), the reaction was stirred for 18 h at rt, followed by evaporation of Pyr. Purification by CC (SiO_2, Hex/Et_2O 3:1) gave a anomeric mixture α/β 1.3:1 of **14** (2.305 g, 6.468 mmol, 86%).

[1] Anomers were separated for analysis, but mixture was used for further synthesis

$\boldsymbol{R_f}$ (SiO_2, Hex/Et_2O 3:1): 0.30 (β–**14**), 0.21 (α–**14**).

IR (film): 3650 *w*, 3063 *w*, 2992 *w*, 2954 *m*, 2932 *m*, 2834 *w*, 1969 *w*, 1915 *w*, 1721 *s*, 1602 *m*, 1584 *w*, 1491 *w*, 1451 *m*, 1375 *m*, 1315 *m*, 1271 *s*, 1210 *m*, 1176 *m*, 1113 *s*, 1069 *s*, 1026 *m*, 936 *w*, 858 *w*, 737 *m*, 711 *s*, 687 *m*.

^{1}H-NMR (600 MHz, $CDCl_3$): (β–**14**): 8.09 (*dd*, $^3J = 8.3$, $^4J = 1.2$, H_{ortho}(Bz)); 8.03 (*dd*, $^3J = 8.3$, $^4J = 1.2$, H_{ortho}(Bz); 7.58 (*dt*, $^3J = 7.5$, $^4J = 1.3$, H_{para}(Bz); 7.55 (*dt*, $^3J = 7.5$, $^4J = 1.3$, H_{para}(Bz); 7.46–7.41 (*m*, H_{meta}(Bz); 5.64–5.61 (*m*, HC(4)); 5.24 (*dd*, $^3J = 5.4$, 2.2, HC(1)); 4.58 (*dd*, $^2J = 10.4$, $^3J = 5.1$, HC(5)); 4.56–4.54 (*m*, HC(3)); 4.50 (*dd*, $^2J = 10.4$, $^3J = 5.3$, HC(5)); 3.37 (*s*, H_3CO); 2.58 (*ddd*, $^2J = 14.1$, $^3J = 7.3$, 2.2, HC(2)); 2.36 (*ddd*, *dt-like*, $^2J = 14.1$, $^3J = 5.4$, HC(2)).
(α–**14**): 8.05 (*dd*, $^3J = 7.1$, $^4J = 1.4$, H_{ortho}(Bz)); 8.03 (*dd*, $^3J = 7.1$, $^4J = 1.4$, H_{ortho}(Bz); 7.60–7.52 (*m*, H_{para}(Bz)); 7.48–7.39 (*m*, H_{meta}(Bz)); 5.46–4.42 (*m*, HC(4); 5.20 (*dd*, $^3J = 5.0$, 2.2, HC(1)); 4.68–4.52 (*m*, HC(3), H_2C(5)); 3.43 (*s*, H_3CO); 2.58 (*ddd*, $^2J = 14.6$, $^3J = 8.2$, 5.0, HC(2)); 2.21 (*ddd*, $^2J = 14.6$, $^3J = 2.2$, 0.8, HC(2)).

^{13}C-NMR (75.5 MHz, $CDCl_3$): (β–**14**): 166.3, 166.1 (CO (Bz)), 133.3, 133.0 (C_{para}(Bz)); 130.0, 129.8, 129.7 (C_{ipso}, C_{ortho}(Bz)); 128.5, 128.4 (C_{meta}(Bz)); 105.6 (C(1)); 81.9 (C(4)); 75.6 (C(3)); 65.3 (C(5)); 55.2 (CH_3); 39.3 (C(2)).
(α–**14**): 166.4, 166.2 (CO (Bz)), 133.1, 133.1 (C_{para}(Bz)); 129.9, 129.8, 129.7 (C_{ipso}, C_{ortho}(Bz)); 128.4 (C_{meta}(Bz)); 105.1 (C(1)); 80.9 (C(4)); 74.8 (C(3)); 64.5 (C(5)); 55.1 (CH_3); 39.3 (C(2)).

ESI-MS: 379.1 (100, $[M+Na]^+$).

6.3 1-(3',5'-di-*O*-Benzyl-2'-deoxy-α/β-D-ribofurano-1-yl)-propan-2-one (10)

Freshly distilled $SnCl_4$ (0.50 ml, 1.111 g, 4.263 mmol, 1.4 eq.) was added under Ar dropwise at 0 °C to a well stirred solution of **8** (1.000 g, 3.045 mmol) and silylenolether[180] (**9**) (0.476 g, 3.654 mmol, 1.2 eq.) in 40 ml dry MeCN. After stirring for 45 min and allow the solution to warm up to rt, MeCN was evaporated. The residue was dissolved in DCM and extracted with *sat.* $NaHCO_3$ solution, dried with $MgSO_4$, and purified by CC (SiO_2, 30 g, Hex/Et_2O 1:1). 1.068 g **10** (3.015 mmol, 99%) of a α/β mixture was optained as a yellowish oil. If starting with α–**10** a α/β 5:1 and if starting with a β–**10** a α/β 2.7:1 mixture was observed. By treating the product with $Zn(OAc)_2$ (6.614 g, 30.132 mmol, 10 eq.) and NaOMe (1.628 g, 30.132 mmol, 10 eq.) in 50 ml MeOH for 2 d at rt the anomeric ratio could be shifted, in both cases, to α/β 1:2.[183]

$\boldsymbol{R_f}$ (SiO_2, Hex/Et_2O 1:1): 0.21.

IR (film): 3407*w*, 3063*m*, 3030*m*, 2892*s*, 2863*s*, 1955*w*, 1877*w*, 1812*w*, 1713*s*, 1605*w*, 1586*w*, 1496*m*, 1453*s*, 1359*s*, 1309*m*, 1242*m*, 1205*m*, 1162*m*, 1098*s*, 1028*s*, 911*m*, 737*s*, 698*s*, 606*w*.

^{1}H-NMR (600 MHz, $CDCl_3$): 7.35–7.25 (*m*, arom. H); 4.56–4.46 (*m*, benz. H, HC(1'), α/β-**10**); 4.20 (*ddd*, *q-like*, ^3J = 4.4, HC(4'), α-**10**); 4.11 (*ddd*, *dt-like*, ^3J = 4.9, 2.6, HC(4'), β-**10**); 4.08 (*ddd*, *dt-like*, ^3J = 7.0, 3.6, HC(3'), α-**10**); 4.02 (*ddd*, *dt-like*, ^3J = 6.3, 1.9, HC(3'), β-**10**); 3.53–3.42 (*m*, H_2C(5'), α/β -**10**); 2.93 (*dd*, ^2J = 16.4, ^3J = 6.8, HC(1), α-**10**); 2.78 (*dd*, ^2J = 16.0, ^3J = 6.9, HC(1), β-**10**); 2.65 (*dd*, ^2J = 16.4, ^3J = 6.5, HC(1), α-**10**); 2.58 (*dd*, ^2J = 16.0, ^3J = 5.8, HC(1), β-**10**); 2.37 (*dt*, ^2J = 13.1, ^3J = 6.7, HC(2'), α-**10**); 2.2 (*ddd*, ^2J = 13.1, ^3J = 5.2, 1.3, HC(2'), β-**10**); 2.17 (*s*, H_3C, β-**10**); 2.16 (*s*, H_3C, α-**10**); 1.75 (*ddd*, ^2J = 13.1, ^3J = 5.6, 4.2, HC(2'), α-**10**); 1.61

(*ddd*, ^2J = 13.1, ^3J = 10.2, 6.3, HC(2'), β-**10**).

^{13}C-NMR (75.5 MHz, $CDCl_3$): 207.2 (CO, α-**10**); 206.7 (CO, β-**10**); 138.1 (arom. $C_{quart.}$, α/β-**10**); 128.3, 128.2, 125.7 (arom. C, α/β-**10**); 83.5, 80.9 (C(3'), C(4'), β-**10**); 82.4, 80.8 (C(3'), C(4'), α-**10**); 74.8 (C(1'), α-**10**); 74.6 (C(1'), β-**10**); 73.3 (C(5'), α/β-**10**); 71.4, 70.9, 70.8, 70.7 (benz. CH_2, α/β-**10**); 49.8 (C(2), α-**10**); 49.2 (C(2), β-**10**); 38.1 (C(2'), β-**10**); 37.7 (C(2'), α-**10**); 30.7 (CH_3, α-**10**); 30.5 (CH_3, β-**10**).

ESI-MS: 377.2 (100, $[M+Na]^+$).

6.4 1-(3',5'-di-*O*-Benzoyl-2'-deoxy-α/β-D-ribofurano-1-yl)-propan-2-one (15)

Freshly distilled $SnCl_4$ (0.06 ml, 0.123 g, 0.471 mmol, 1.4 eq.) was added under Ar dropwise at 0 °C to a well stirred solution of **14** (0.120 g, 0.337 mmol, α/β 1.3:1) and silylenolether[180] **9** (0.053 g, 0.404 mmol, 1.2 eq.) in 10 ml dry MeCN. After stirring for 45 min and allow the solution to warm up to rt, MeCN was evaporated. The residue was dissolved in DCM and extracted with *sat.* $NaHCO_3$ solution, dried with Na_2SO_4 and purified by CC (SiO_2, 5 g, Hex/Et_2O 1:2). 0.037 g **15** (0.097 mmol, 29%, α/β 2:1) was optained as a yellowish oil.

$\boldsymbol{R_f}$ (SiO_2, Hex/Et_2O 1:2): 0.27.

IR (film): 3045 *w*, 2966 *m*, 2930 *m*, 1961 *w*, 1912 *w*, 1729 *s*, 1717 *s*, 1613 *m*, 1591 *w*, 1444 *m*, 1372 *w*, 1343 *w*, 1332 *m*, 1282 *m*, 1254 *w*, 1209 *w*, 1198 *w*, 1124 *s br*, 965 *m*, 854 *w*, 831 *w*, 721 *s*, 699 *s*, 675 *m*.

^{1}H-NMR (600 MHz, $CDCl_3$): 8.06–8.01 (*m*, H_{ortho}(Bz)); 7.62–7.53 (*m*, H_{para}(Bz)); 7.48–7.27 (*m*, H_{meta}(Bz)); 5.54–5.28 (*m*, HC(4')); 4.72 (*q*, ^3J = 6.6, HC(1'), α-**15**); 4.65–4.36 (*m*, H_2C(5'), HC(3'), HC(1') (β-**15**)); 2.99 (*dd*, ^2J = 16.4, ^3J = 7.0, HC(1), α-**15**); 2.88 (*dd*, ^2J = 16.3, ^3J = 6.9, HC(1), β-**15**); 2.79–2.74 (*m*, H_αC(2'), α/β-**15**); 2.73 (*dd*, ^2J = 16.4, ^3J = 6.3, HC(1), α-**15**); 2.66 (*dd*, ^2J = 16.3, ^3J = 5.5, HC(1), β-**15**); 1.98 (*ddd*, ^2J = 14.0, ^3J = 5.6, 3.7, H_βC(2'), α-**15**); 1.95 (*ddd*, ^2J = 14.0, ^3J = 5.6, 3.6, H_βC(2'), β-**15**).

^{13}C-NMR (75.5 MHz, $CDCl_3$): 206.3, 206.2 (C(2)); 166.1, 165.9 (CO(Bz)); 133.3, 133.0 (C_{para}(Bz)); 129.8, 129.7, 129.6, 129.5 (C_{ipso}, C_{ortho}(Bz)); 128.4, 128.3 (C_{meta}(Bz)); 82.4, 81.4 (C(4')); 76.8, 76.6 (C(1')); 75.0, 74.9 (C(3')); 64.6, 64.5 (C(5')); 49.6, 48.7 (C(1)); 38.5, 37.7 (C(2')); 30.7, 30.5 (C(3)).

ESI-MS: 405.4 (100, $[M+Na]^+$).

6.5 2-Amino-4-((3',5'-di-*O*-benzyl-2'-deoxy-β-D-ribofurano-1-yl)methyl)pyrimidine (12β)

According to the literature,[177,181] ^{t}BuOCH$(NMe_2)_2$ (0.369 g, 2.116 mmol, 1.5 eq.) was synthesized, added to a solution of **10** (0.500 g, 1.411 mmol) in 50 ml Toluene and stirred for 20 h at 70 °C. After solvent evaporation the residue was dissolved in 16 ml EtOH, followed by addition of NaOEt (0.192 g, 2.822 mmol, 2 eq.) and stirred for 4 h at 60 °C.[2] After cooling to 40 °C Guanidinium sulfate (1.220 g, 5.643 mmol, 4 eq.) was added and reheated to 70 °C for 18 h. The solvent was evaporated and purification by CC (SiO_2, 20 g, Hex/EE 1:8) gave 0.418 g of a yellowish oil, which was a mixture of desired product anomers **12**α/β (α/β 1:2) containing 12% of 2-Amino-4-methyl-5-

[2]The Enaminoketone intermediate was not isolated but detected by ^{13}C-NMR and MS (see analysis).

(3,5-di-*O*-benzyl-2-deoxy-α/β-pentofuranosyl)-pyrimidine. This mélange was separated by peparative HPLC (Hex/MeOH/EtOH 97:2:1) yielding 0.273 g **12**β (0.923 mmol, 48%).

^{13}C-NMR of Enaminoketon (75.5 MHz, $CDCl_3$): 108.1, 107.4 (C(1)H)

ESI-MS of Enaminoketon: 432.3 (100, $[M+Na]^+$).

$\boldsymbol{R_f}$ (SiO_2, Hex/EE 1:8): 0.23.

EA: calc.: %C: 71.09, %H: 6.71, %O: 10.36; found: %C: 70.54, %H: 6.77, %O: 9.81.

IR (film): 3325*s*, 3179*m*, 3087*m*, 3062*m*, 3029*m*, 2923*m*, 2862*s*, 1955*w*, 1887*w*, 1815*w*, 1629*s*, 1576*s*, 1495*m*, 1454*s*, 1362*m*, 1341*m*, 1266*m*, 1204*m*, 1094*s*, 1060*s*, 1027*m*, 912*w*, 812*w*, 736*s*, 697*s*, 604*w*.

^{1}H-NMR (500 MHz, $CDCl_3$): 8.14 (*d*, $^3J = 5.1$, HC(6); 7.37–7.24 (*m*, arom. H); 6.55 (*d*, $^3J = 5.1$, HC(5)); 5.06 (*s br*, H_2N); 4.54 (*s*, $H_2COC(5')$); 4.55–4.46 (*m*, HC(1')); 4.48 (*s*, $H_2COC(3')$); 4.13 (*ddd*, *dt-like*, $^3J = 5.2$, 2.6, HC(4')); 4.02 (*ddd*, *dt-like*, $^3J = 6.7$, 2.0, HC(3')); 3.52 (*dd*, $^2J = 10.1$, $^3J = 4.7$, HC(5')); 3.42 (*dd*, $^2J = 10.1$, $^3J = 5.4$, HC(5')); 2.89 (*dd*, $^2J = 13.9$, $^3J = 6.9$, HCC(4)); 2.75 (*dd*, $^2J = 13.9$, $^3J = 5.7$, HCC(4)); 2.10 (*ddd*, $^2J = 13.2$, $^3J = 5.2$, 1.5, $H_\alpha C(2')$); 1.72 (*ddd*, $^2J = 13.2$, $^3J = 10.2$, 6.3, $H_\beta C(2')$).

^{13}C-NMR (125.8 MHz, $CDCl_3$): 168.6 (C(4)); 162.9 (C(2)); 158.0 (C(6)); 138.2, 138.1 (arom. $C_{quart.}$); 128.5, 128.4, 128.2, 127.8, 127.7, 127.6 (arom. C); 111.6 (C(5)); 83.6 (C(4')); 81.0 (C(3')); 77.4 (C(1')); 73.4 (*C*OC(5')); 71.0 (*C*OC(3')); 70.9 (C(5')); 43.4 (*C*C(4)); 38.0 (C(2')).

ESI-MS: 428.2 (100, $[M+Na]^+$).

HPLC (Hex/MeOH/EtOH 97:2:1): t_R = 22.1 min, k' = 6.94, α_{21} = 1.10, α_{32} = 1.08.

6.6 2-Amino-4-((3',5'-di-*O*-benzyl-2'-deoxy-α-D-ribofurano-1-yl)methyl)pyrimidine (**12**α)

12α was obtained as a side product of the synthesis of **12**β and isolated by HPLC purification, as described in *Section 6.5*.

$\boldsymbol{R_f}$ (SiO_2, Hex/EE 1:8): 0.23.

^{1}H-NMR (600 MHz, $CDCl_3$): 8.16 (*d*, ^3J = 5.1, HC(6); 7.36–7.24 (*m*, arom. H); 6.56 (*d*, ^3J = 5.1, HC(5)); 4.95 (*s br*, H_2N); 4.58–4.46 (*m*, HC(1'), H_2COC(3'), H_2COC(5')); 4.24 (*ddd*, *q-like*, ^3J = 4.4, HC(4')); 4.11 (*ddd*, ^3J = 6.8, 4.4, 3.7, HC(3')); 3.52 (*dd*, ^2J = 10.2, ^3J = 4.7, HC(5')); 3.47 (*dd*, ^2J = 10.2, ^3J = 4.9, HC(5')); 3.02 (*dd*, ^2J = 13.7, ^3J = 7.5, HCC(4)); 2.79 (*dd*, ^2J = 13.7, ^3J = 5.9, HCC(4)); 2.29 (*ddd*, *dt-like*, ^2J = 13.0, ^3J = 6.8, H_αC(2')); 1.86 (*ddd*, ^2J = 13.0, ^3J = 6.1, 4.4, H_βC(2')).

^{13}C-NMR (125.8 MHz, $CDCl_3$): 168.5 (C(4)); 163.1 (C(2)); 157.9 (C(6)); 138.2, 138.1 (arom. $C_{quart.}$); 128.5, 128.3, 127.7, 127.6(arom. C); 111.5 (C(5)); 83.6 (C(4')); 81.4 (C(3')); 77.4 (C(1')); 73.3 (*C*OC(5')); 72.1 (*C*OC(3')); 70.3 (C(5')); 43.0 (*C*C(4)); 36.8 (C(2')).

ESI-MS: 428.2 (100, $[M+Na]^+$).

HPLC (Hex/MeOH/EtOH 97:2:1): t_R = 23.5 min, k' = 7.46, α_{32} = 1.08, α_{43} = 1.17.

6.7 2-Amino-4-methyl-5-(3',5'-di-*O*-benzyl-2'-deoxy-β-D-ribofurano-1-yl)pyrimidine (**13**β)

13β was obtained as a side product of the synthesis of **12**β and isolated by HPLC purification, as described *Section 6.5*.

$\boldsymbol{R_f}$ (SiO_2, Hex/EE 1:8): 0.23.

IR (KBr): 3339*s*, 3182*s*, 2906*m*, 2867*m*, 2837*m*, 1660*s*, 1594*s*, 1557*s*, 1488*s*, 1453*m*, 1364*s*, 1357*s*, 1313*m*, 1218*m*, 1187*m*, 1158*m*, 1102*s*, 1057*s*, 800*m*, 743*s*, 698*s*, 606*m*, 554*m*.

^{1}H-NMR (600 MHz, $CDCl_3$): 8.30 (*s*, HC(6)); 7.37–7.25 (*m*, arom. H); 5.16 (*dd*, ^3J = 10.8, 5.2, HC(1')); 5.06 (*s br*, H_2N); 4.59 (*d*, ^2J = 12.2, HCOC(5')); 4.55 (*s br*, $H_2COC(3')$); 4.52 (*d*, ^2J = 12.2, HCOC(5')); 4.24 (*ddd*, *dt-like*, ^3J = 4.8, 2.4, HC(4')); 4.18–4.15 (*m*, HC(3')); 3.63 (*dd*, ^2J = 10.1, ^3J = 4.5, HC(5')); 3.55 (*dd*, ^2J = 10.1, ^3J = 5.1, HC(5')); 2.35 (*s*, H_3C); 2.31 (*ddd*, ^2J = 13.1, ^3J = 5.2, 1.2, $H_\alpha C(2')$); 1.87 (*ddd*, ^2J = 13.1, ^3J = 10.8, 6.1, $H_\beta C(2')$).

^{13}C-NMR (75.5 MHz, $CDCl_3$): 165.6 (C(2)); 162.0 (C(6)); 156.0 (C(4)); 137.9 (C_{tert}(Bn)); 128.4, 128.3, 127.7, 127.6, 127.5 (arom. C); 122.2 (C(5)); 83.6 (C(4')); 81.2 (C(1')); 75.4 (C(3')); 73.4 (C(5')); 71.1, 70.8 (benz. C); 39.2 (C(2')); 21.6 (CH_3).

ESI-MS: 428.2 (100, $[M+Na]^+$).

HPLC (Hex/MeOH/EtOH 97:2:1): t_R = 20.4 min, k' = 6.32, α_{21} = 1.10.

6.8 2-Amino-4-methyl-5-(3',5'-di-*O*-benzyl-2'-deoxy-α-D-ribofurano-1-yl)pyrimidine (**13**α)

13α was obtained as a side product of the synthesis of **12**β and isolated by HPLC purification, as described *Section 6.5*.

$\boldsymbol{R_f}$ (SiO_2, Hex/EE 1:8): 0.23.

^{1}H-NMR (600 MHz, $CDCl_3$): 8.37 (*s*, HC(6)); 7.37–7.25 (*m*, arom. H); 5.14 (*dd*, *br t-like*, ^3J = 7.4, HC(1')); 5.01 (*s br*, H_2N); 4.60 (*d*, ^2J = 12.2, HCOC(5')); 4.56 (*d*, ^2J = 12.2, HCOC(5')); 4.50 (*s*, $H_2COC(3')$); 4.34 (*ddd*, *q-like*, ^3J = 4.3, HC(4')); 4.26 (*ddd*, ^3J = 6.7, 5.6, 4.2, HC(3')); 3.61 (*dd*, ^2J = 10.3, ^3J = 4.3, HC(5')); 3.59 (*dd*, ^2J = 10.3, ^3J = 4.5, HC(5')); 2.58 (*ddd*, *dt-like*, ^2J = 12.8, ^3J = 6.8, $H_\beta C(2')$); 2.35 (*s*, H_3C); 1.98 (*ddd*, ^2J = 12.8, ^3J = 8.1, 5.6, $H_\alpha C(2')$).

ESI-MS: 428.2 (100, $[M+Na]^+$).

HPLC (Hex/MeOH/EtOH 97:2:1): t_R = 27.0 min, k' = 8.72, α_{43} = 1.17.

6.9 2-Amino-4-((2'-deoxy-β-D-ribofurano-1-yl)methyl)pyrimidine (**1**)

BBr_3 (2.96 ml, 1M in DCM, 2.959 mmol, 6 eq.) was added at 0 °C to a solution of **12**β (0.200 g, 0.493 mmol) in 15 ml *abs.* DCM. The reaction was quenched with 10 ml H_2O after 3 min, neutralized with *sat.* $NaHCO_3$ solution and extracted with DCM. The water phase was evaporated and purified by CC (SiO_2 , 3 g, EE/MeOH 10:1), yielding 0.110 g (0.488 mmol, 99%) of **1** as a white solide.

$\boldsymbol{R_f}$ (SiO_2, EE/MeOH 10:1): 0.10.

UV (H_2O): λ_{max} = 291.8, log ϵ = 3.581; λ_{max} = 225.7, log ϵ = 4.057; λ_{max} = 185.8, log ϵ = 4.013; λ_{min} = 249.8, log ϵ = 2.776; λ_{min} = 206.6, log ϵ = 3.721.

IR (KBr): 3342*s*, 3223*s*, 2922*s*, 2875*s*, 2136*w*, 1632*s*, 1582*s*, 1567*s*, 1468*s*, 1340*m*, 1269*m*, 1218*m*, 1168*m*, 1096*s*, 1048*s*, 909*w*, 876*w*, 804*w*, 781*w*, 714*w*.

^{1}H-NMR (400 MHz, MeOD): 8.13 (*d*, ^3J = 5.1, HC(6)); 6.64 (*d*, ^3J = 5.1, HC(5)); 4.52–4.45 (*m*, HC(1')); 4.21 (*ddd*, *dt-like*, ^3J = 6.2, 2.9, HC(4')); 3.77 (*ddd*, *dt-like*, ^3J = 4.8, 2.9, HC(3')); 3.55 (*dd*, ^2J = 11.7, ^3J = 4.6, HC(5')); 3.51 (*dd*, ^2J = 11.7, ^3J = 5.2, HC(5')); 2.83 (*dd*, ^2J = 13.7, ^3J = 7.3, HCC(4)); 2.78 (*dd*, ^2J = 13.7, ^3J = 5.6, HCC(4)); 1.95 (*ddd*, ^2J = 13.1, ^3J = 5.5, 2.2, HC(2')); 1.82 (*ddd*, ^2J = 13.1, ^3J = 9.7, 6.2, HC(2')).

^{13}C-NMR (75.5 MHz, MeOD): 171.9 (C(4)); 165.7 (C(2)); 160.2 (C(6)); 113.1 (C(5)); 90.1 (C(4')); 79.9 (C(3')); 75.2 (C(1')); 65.2 (C(5')); 45.6 (C(2')); 43.0 (*C*C(4)).

ESI-MS: 226.0 (100, $[M+H]^+$); 248.0 (77, $[M+Na]^+$).

6.10 *N*-(4-((2'-deoxy-β-D-ribofurano-1-yl)-methyl)pyrimidin-2-yl)benzamide (18)

A solution of **1** (0.100 g, 0.444 mmol) and TMSCl (0.28 ml, 2.220 mmol, 5 eq.) in 10 ml Pyr was stirred for 12 h at 0 °C, followed by addition of BzCl (0.26 ml, 2.220 mmol, 5 eq.) and stirring for 12 h at rt.[187] 0.6 ml H_2O and 2 ml of *aq.* 25% NH_3 solution was added to quench the reaction. The

solvent was evaporated and the residue was extracted with Et_2O and H_2O. The aqueous phase was evaporated and dried under HV. The obtained white solid was purified by CC (SiO_2, 2.5 g, EE/MeOH 10:1), giving 0.080 g (0.244 mmol, 55%) of **18**.

$\boldsymbol{R_f}$ (SiO_2, EE/MeOH 10:1): 0.12.

IR (KBr): 3406*s*, 2935*m*, 2526*w*, 2126*w*, 1686*s*, 1600*s*, 1529*s*, 1493*m*, 1441*s*, 1404*s*, 1341*m*, 1269*s*, 1192*w*, 1099*m*, 1076*m*, 1051*m*, 1002*w*, 905*w*, 796*w*, 713*m*.

^{1}H-NMR (400 MHz, MeOD): 8.53 (*d*, $^3J = 5.1$, HC(6)); 8.02–7.97 (*m*, arom. H_{meta}); 7.62–7.43 (*m*, arom. $H_{ortho,para}$); 7.18 (*d*, $^3J = 5.1$, HC(5)); 4.61–4.55 (*m*, HC(1')); 4.22 (*ddd*, *dt-like*, $^3J = 6.2$, 2.5, HC(4')); 3.79 (*ddd*, *dt-like*, $^3J = 4.9$, 2.8, HC(3')); 3.55 (*dd*, $^2J = 11.7$, $^3J = 4.7$, HC(5')); 3.52 (*dd*, $^2J = 11.7$, $^3J = 5.0$, HC(5')); 3.00 (*d*, $^3J = 6.3$, H_2CC(4)); 1.99 (*ddd*, $^2J = 13.1$, $^3J = 5.5$, 2.2, H_αC(2')); 1.86 (*ddd*, $^2J = 13.1$, $^3J = 9.8$, 6.1, H_βC(2')).

^{13}C-NMR (75.5 MHz, MeOD): 171.2 (C(4)); 168.4 (CO); 159.0, 158.9 (C(2), C(6)); 135.6, 133.6, 129.7, 129.1 (arom. C); 118.5 (C(5)); 89.0 (C(4')); 78.4 (C(3')); 73.9 (C(1')); 63.9 (C(5')); 44.2 (C(2')); 41.7 (*C*C(4)).

ESI-MS: 330.3 (6, $[M+H]^+$); 352.2 (100, $[M+Na]^+$).

6.11 *N*-(4-((5'-*O*-(4",4"'-dimethoxytrityl)-2'-deoxy-β-D-ribofurano-1-yl)methyl)pyrimidin-2-yl)benzamide (19)

Compound **18** (0.057 g, 0.173 mmol), DMTCl (0.076 g, 0.225 mmol, 1.3 eq.) and $(Bu_4N)ClO_4$ (0.071 g, 0.208 mmol, 1.2 eq.) were dried overnight under HV. 3 ml Pyr was added and stirred at rt for 16 h. After evaporation of Pyr the residue was dissolved in DCM and extracted with *sat.* $NaHCO_3$ and brine. The organic phase was dried with Na_2SO_4, filtrated and evaporated. After CC (SiO_2, 3 g, EE (0.5% Et_3N)), **19** (0.085 g, 0.135 mmol, 78%) was obtained as a white foam.

$\boldsymbol{R_f}$ (SiO_2, EE (0.5% Et_3N)): 0.18.

^{1}H-NMR (300 MHz, DCM): 8.44 (*d*, ^3J = 5.0, HC(6)); 7.86 (*d*, ^3J = 7.1, 2 benz. H); 7.53 (*t*, ^3J = 7.4, 1 benz. H); 7.46–7.17 (*m*, 11 arom. H); 6.98 (*d*, ^3J = 5.0, HC(5)); 6.80 (*d*, ^3J = 8.4, 4 arom. H); 4.58–4.54 (*m*, HC(1)); 4.28 (*ddd*, *dt-like*, ^3J = 6.0, 2.3, HC(4)); 3.89 (*ddd*, *dt-like*, ^3J = 4.8, 2.7, HC(3)); 3.74 (*s*, H_3CO); 3.72–3.13 (*m*, $H_2C(5)$); 3.10 (*dd*, ^2J = 9.8, ^3J = 4.8, HCC(4)); 3.05 (*dd*, ^2J = 9.8, ^3J = 5.1, HCC(4)); 2.00 (*ddd*, ^2J = 13.1, ^3J = 5.5, 2.1, $H_\alpha C(2)$); 1.85 (*ddd*, ^2J = 13.1, ^3J = 9.8, 6.0, $H_\beta C(2)$).

^{13}C-NMR (75.5 MHz, DCM): 169.8 (C(4)); 165.3 (CON); 158.9 ($C_{arom.}$OMe); 158.2 (C(6)); 158.0 (C(2)); 145.5 (C_{ipso}(Ph)); 136.4 (C_{ipso}(PhOMe)); 134.9 (C_{ipso}(Bz)); 132.6 (C_{para}(Bz)); 130.4 (C_{ortho}(PhOMe)); 129.1, 128.5, 128.2, 127.9 (C_{ortho}(Ph), C_{meta}(Ph), C_{ortho}(Bz), C_{meta}(Bz)); 127.1 (C_{para}(Ph)); 117.3 (C(5)); 113.4 (C_{meta}(PhOMe)); 86.5 (C(4')); 86.4 (C_{tert}(Tr)); 77.2 (C(3')); 74.3 (C(1')); 65.0 (C(5')); 55.6 (CH_3O); 43.9 (*C*C(4)); 41.2 (C(2')).

ESI-MS: 654.3 (100, $[M+Na]^+$).

HR-ESI-MS: calc. for $C_{38}H_{37}N_3O_6Na$ ($[M+Na]^+$) 654.2578; found 654.2578.

6.12 (1-((N$^{2'}$-Benzoyl(2-aminopyrimidin-4-yl))methyl)-5-*O*-(4,4-dimethoxytrityl)-2-deoxy-β-D-ribofurano-1-yl)-β-cyanoethyl-*N*,*N*-diisopropyl phosphoramidite (20)

$NEt(^iPr)_2$ (0.09 ml, 0.507 mmol, 4 eq.) and $PCl(N(^iPr)_2)(OEtCN)$ (0.03 ml, 0.158 mmol, 1.25 eq.) were added dropwise under Ar to a solution of **19** (0.080 g, 0.127 mmol) in 2 ml *abs.* DCM at rt. After 2 h, the reaction mixture was diluted with 5 ml of EE (0.5% Et_3N) and extracted with 10% Na_2CO_3 solution and brine, dried with $MgSO_4$, evaporated and purified by CC (SiO_2, 3g, EE/Hex 2:1 (0.1% Et_3N)), yielding **20** as a colourless oil (0.068 g, 0.0818 mmol, 64%).

$\boldsymbol{R_f}$ (SiO_2, EE (0.1% Et_3N)): 0.21, 0.15.

^{1}H-NMR (300 MHz, DCM): 8.72, 8.69 (2 *s br*, NH); 8.50, 8.49 (2 *d*, ^3J = 5.0, HC(6')); 7.88–7.85 (*m*, 2 benz. H_{ortho}); 7.56 (*t br*, ^3J = 7.4, benz. H_{para}); 7.49–7.18 (*m*, 11 arom. H); 7.05, 7.04 (2 *d*, ^3J = 5.0, HC(5')); 6.83, 6.82 (2 *d*, ^3J = 8.8, HC_{meta}(PhOMe)); 4.60–4.53 (*m*, HC(1)); 4.46–4.39 (*m*, HC(4)); 4.08–4.02 (*m*, HC(3)); 3.76, 3.75 (2 *s*, H_3CO), 3.81–3.51 (*m*, $H_2C(5)$, OCH_2CH_2CN); 3.16–3.00 (*m*, $NCHMe_2$, $H_2CC(4')$); 2.57, 2.44 (2 *t*, ^3J = 6.3, OCH_2CH_2CN); 2.25–2.10 (*m*, $H_\alpha C(2)$); 1.91 (*ddd*, ^2J = 13.0, ^3J = 9.8, 6.1, $H_\beta C(2)$).

^{13}C-NMR (75.5 MHz, DCM): 169.7 (C(4')); 165.0 (CON); 159.0 (COMe);

158.3, 158.0 (C(2'), C(6')); 145.5 (C_{ipso}(Ph)); 136.5 ($C_{ipso,para}$(PhOMe)); 135.1 (C_{ipso}(Bz)); 132.5 (C_{para}(Bz)); 130.5, 129.1, 128.6, 128.2, 127.9 ($C_{ortho,meta}$(Ph), $C_{ortho,meta}$(Bz), C_{ortho}(PhOMe)); 127.1 (C_{para}(Ph)); 118.2, 118.1 (CN); 117.3 (C(5')); 113.4 (C_{meta}(PhOMe)); 86.4 (C_{tert}(Tr)); 86.0, 85.8 (C(4)); 77.6, 77.5 (C(1)); 75.7, 75.5 (C(3)); 64.6 (C(5)); 58.8, 58.6 (CH$_2$CH$_2$CN); 55.6 (CH$_3$O); 44.0 (CC(4')); 43.6, 43.5 (CH$_3$$C$HN); 40.4 (C(2)); 24.7 ($CH_3$CHN); 21.3, 20.7 (CH$_2$$CH_2$CN).

ESI-MS: 854.4 (100, $[M+Na]^+$).

6.13 2-Iodo-4-methyl-5-(3',5'-di-*O*-benzyl-2'-deoxy-α/β-D-ribofurano-1-yl)pyrimidine (**22**α/β)

An anomeric mixture of **13**α/β (0.069 g, 0.170 mmol), Isopentyl nitrite (0.07 ml, 0.510 mmol, 3 eq.), CH_2I_2 (0.07 ml, 0.851 mmol, 5 eq.), and CuI (32.4 mg, 0.170 mmol, 1 eq.) in 2 ml dry THF was refluxed under Ar for 1.5 h, diluted with DCM and extracted with *sat.* NaCl solution. The organic phase was dried with Na_2SO_4, filtrated and evaporated. CC (SiO_2, 4 g, DCM/EE 20:1) gave **22** (0.017 g, 0.0329 mmol, 19.4%) as a brownish resin.

$\boldsymbol{R_f}$ (SiO_2, DCM/EE 20:1): 0.18.

^{1}H-NMR (300 MHz, $CDCl_3$): 8.50, 8.49 (2 *s*, HC(6), α/β-**22**)); 7.38–7.26 (*m*, arom. H); 5.24–5.17 (*m*, HC(1')); 4.61–4.54 (*m*, benz. H); 4.30–4.18 (*m*, HC(3'), HC(4')); 3.64–3.54 (*m*, H_2C(5')); 2.69–2.60 (*m*, HC(2'), α/β-**22**); 2.43 (*s*, H_3C, α-**22**)); 2.39 (*s*, H_3C, β-**22**)); 1.93 (*ddd*, ^2J = 13.0, ^3J = 6.2, 4.0, HC(2'), β-**22**); 1.81 (*ddd*, ^2J = 13.0, ^3J = 10.7, 5.9, HC(2'), α-**22**).

^{13}C-NMR (75.5 MHz, $CDCl_3$): 166.2, 164.5 (C(4)); 156.1, 155.9 (C(6));

137.7 (C_{tert}(Bn)); 132.8, 131.3 (C(5)); 128.4, 128.3, 127.8, 127.7, 127.6, 127.5, 127.5 (C(2), arom. C); 84.0, 83.6 (C(4')); 81.1, 80.6 (C(1')); 75.3, 75.0 (C(3')); 73.5 (CH$_2$OC(3')); 71.6, 71.3 (CH$_2$OC(5')); 70.8, 70.5 (C(5')); 39.4, 38.8 (C(2')); 21.5 (CH_3).

ESI-MS: 539.1 (100, $[M+Na]^+$).

6.14 2-*O*-benzyl-4-methyl-5-(3',5'-di-*O*-benzyl-2'-deoxy-α/β-D-ribofurano-1-yl)pyrimidine (**23**α/β)

NaOH (9.8 mg, 0.097 mmol, 5 eq.) and BnOH (0.08 ml, 0.775 mmol, 40 eq.) were added to a solution of **22** (10 mg, 0.0194 mmol) in 1 ml MeCN.[189] After stirring for 3.5 h at 85 °C under Ar, the solvent was evaporated and the residue was purified by CC (SiO_2, 1 g, EE/DCM 1:1), yielding 9.5 mg **23** (0.0192 mmol, 99%) as a colorless oil.

$\boldsymbol{R_f}$ (SiO_2, EE/DCM 1:1): 0.18.

^{1}H-NMR (300 MHz, $CDCl_3$): 8.57, 8.52 (2 *s*, HC(6), α/β-**23**)); 7.60–7.26 (*m*, arom. H); 5.50, 5.42 (2 *s*, H_2COC(2), α/β-**23**)); 5.24–5.19 (*m*, HC(1')); 4.58, 4.57, 4.56, 4.55 (4 *s*, benz. H); 4.29–4.25 (*m*, HC(4')); 4.19–4.17 (*m*, HC(3')); 3.66–3.54 (*m*, H_2C(5')); 2.69–2.58 (*m*, HC(2'), β-**23**); 2.44 (*s*, H_3C, α-**23**)); 2.41 (*s*, H_3C, β-**23**)); 2.39–2.33 (*m*, HC(2'), α-**23**); 2.03–1.92 (*m*, HC(2'), β-**23**); 1.86 (*ddd*, $^2J = 13.1$, $^3J = 10.8$, 6.0, HC(2'), α-**23**).

^{13}C-NMR (75.5 MHz, $CDCl_3$): 167.8 (C(4)); 164.1 (C(2)); 157.6 (C(6)); 137.8 (C_{tert}(Bn)); 137.1 (C_{tert}CH$_2$OC(2)); 128.4, 128.3, 128.2, 127.9, 127.7, 127.6, 127.6 (arom. C); 126.4 (C(5)); 84.1 (C(4')); 80.9 (C(1')); 75.2 (C(3')); 73.4, 71.2, 70.7, 68.7 (benz. CH_2, C(5')); 39.4 (C(2')); 21.8 (CH_3).

ESI-MS: 497.6 (31, $[M+H]^+$), 519.2 (100, $[M+Na]^+$), 535.1 (9, $[M+K]^+$).

6.15 4-methyl-5-(2'-deoxy-α/β-D-ribofurano-1-yl)pyrimidin-2-one (24α/β)

BBr_3 (96.7 μl, 0.0967 mmol, 6 eq.) was added to a solution of **23** (8 mg, 0.0161 mmol) in 1.3 ml dry CH_2Cl_2 at 0 °C.[223] The reaction was quenched with 0.6 ml H_2O after 3 min and neutralized with *sat.* $NaHCO_3$. After phase separation, the aqueous solution was evaporated and filtrated over SiO_2 (0.5 g, DCM/MeOH 10:1), giving **24** (3.6 mg, 0.0159 mmol, 99%) as a white solid.

$\boldsymbol{R_f}$ (SiO_2, DCM/MeOH 10:1): 0.28.

^{1}H-NMR (300 MHz, MeOD): 8.21 (*s br*, HC(6)); 5.01–5.04 (*m*, HC(1')); 4.39–4.24 (*m*, HC(3'), HC(4')); 3.62–3.56 (*m*, H_2C(5')); 2.63–2.55 (*m*, HC(2'), β-**24**); 2.38–2.31 (*m*, HC(2'), α-**24**); 2.35 (*s*, H_3C, α-**24**)); 2.33 (*s*, H_3C, β-**24**)); 2.20–12 (*m*, HC(2'), β-**24**); 1.94–1.85 (*m*, HC(2'), α-**24**).

ESI-MS: 227.0 (100, $[M+H]^+$), 249.0 (38, $[M+Na]^+$), 265.0 (9, $[M+K]^+$).

6.16 Synthesis of Silyl Protected 2-deoxy-D-ribono-1,4-lactones

The two silyl protected lactones 3,5-Di-*O*-((^{t}butyldimethylsilyl)-2-deoxy-D-ribono-1,4-lactone (**26a**) and 3,5-*O*-((1',1',3',3'-Tetraisopropyl)disiloxandiyl)-2-deoxy-D-ribono-1,4-lactone (**26b**) were synthesized from commercially available 2-Deoxy-D-ribose in good yields, according to the literature.[162,201,202]

6.17 4-((*Z*)-4',6'-Di-*O*-(^{t}butyldimethylsilyl)-2',4',5',6'-tetrahydroxyhex-1'-en-1'-yl)-pyrimidin-2(*1H*)-one (30a)

n-BuLi (0.49 ml 2M in Heptan, 0.972 mmol, 3.5 eq.) was added to freshly distilled NH(^{i}Pr)$_2$ (0.07 ml, 0.972 mmol, 3.5 eq.) in 3 ml Et_2O at 0 °C, stirred for 30 min and admixed to a suspension of 4-Methylpyrimidin-2-one (0.045 g, 0.305 mmol, 1.1 eq.) in 7 ml THF at –78 °C. After 1 h, the suspension was allowed to warm up to 0 °C and **26a** (0.100 g, 0.278 mmol) was added as a solid. The mixture was reacted overnight, quenched with 0.2 ml MeOH and extracted with 1M NH_4Cl and DCM. The organic phase was dried with Na_2SO_4, filtrated and evaporated. Purification by CC (3g SiO_2, EE/DCM 1:1) gave **30a** as a yellow solid (0.033 g, 0.0701 mmol, 25%).

$\boldsymbol{R_f}$ (SiO_2, EE/DCM 1:1): 0.33.

^{1}H-NMR (400 MHz, $CDCl_3$): 12.64, 12.60 (2 *s*, HOC(2')); 7.25 (*s br*, NH); 6.87 (d, ^3J = 7.4, HC(6), (*3H*)-Pyr); 6.87 (dd, ^3J = 7.4, 10.3, HC(6), (*1H*)-Pyr); 5.55, 5.53 (2 *d*, ^3J = 7.4, HC(5), (*1H*)/(*3H*)-Pyr); 5.23 (*s*, HC(1')); 4.22–4.10, 3.68-3.58 (2 *m*, HC(4'), HC(5'), H_2C(6'), HOC(5')); 2.67–2.46 (m, H_2C(3')); 0.89, 0.88, 0.87, 0.84 (4 *s*, H_3C(^{t}Bu)); 0.08, 0.07, 0.06, 0.05 (4 *s*, H_3CSi).

^{13}C-NMR (100.6 MHz, $CDCl_3$): 199.9, 198.6 (C(2')); 151.6, 151.1, 150.4, 150.2 (C(4), C(2)); 135.9, 135.7 (C(6)); 101.6 (C(1'), (*3H*)-Pyr); 95.6, 94.9 (C(5)); 74.9 (C(1'), (*1H*)-Pyr); 71.2 (C(5')); 70.3 (C(4')); 65.3, 63.7 (C(6')); 46.5, 44.4 (C(3')); 25.8, 25.7 (CH_3(^{t}Bu)); 18.2, 18.0 (C(CH_3)$_3$); -4.5, -4.8, -5.5, -5.6 (CH_3Si).

ESI-MS: 471.3 (20, $[M+H]^+$); 493.3 (100, $[M+Na]^+$); 509.3 (11, $[M+K]^+$).

HR-ESI-MS: calc. for $C_{22}H_{42}N_2O_5Si_2Na$ ($[M+Na]^+$) 493.2524; found 493.2526.

6.18 4-((*Z*)-4',6'-*O*-((1",1",3",3"-Tetraisopropyl)disiloxandiyl)-2',4',5',6'-tetrahydroxyhex-1'-en1'-yl)pyrimidin-2(*1H*)-one (30b)

n-BuLi (0.49 ml 2M in Heptan, 0.972 mmol, 3.6 eq.) was added to freshly distilled $NH(^iPr)_2$ (0.07 ml, 0.972 mmol, 3.6 eq.) in 3 ml Et_2O at 0 °C, stirred for 30 min and admixed to a suspension of 4-Methylpyrimidin-2-one (0.045 g, 0.305 mmol, 1.14 eq.) in 7 ml THF at –78 °C. After 1 h, the suspension was allowed to warm up to 0 °C and **26b** (0.100 g, 0.267 mmol) was added as a solid. The mixture was reacted overnight, quenched with 0.2 ml MeOH and extracted with 1M NH_4Cl and DCM. The organic phase was dried with Na_2SO_4, filtrated and evaporated. Purification by CC (3g SiO_2, EE/DCM 1:1) gave **30b** as a yellow solid (0.033 g, 0.0681 mmol, 25%).

$\boldsymbol{R_f}$ (SiO_2, EE/DCM 1:1): 0.25.

^{1}H-NMR (400 MHz, $CDCl_3$): 12.64 (*s*, HOC(2')); 10.17 (*s br*, HN(1)); 6.91 (*dd*, ^3J = 7.4, 5.5, HC(6)); 5.55 (*d*, ^3J = 7.4, HC(5)); 5.27 (*s*, HC(1')); 4.24 (*ddd*, *dt-like*, ^3J = 9.1, 5.3, HC(4')); 4.13 (*dd*, ^2J = 11.7, ^3J = 1.1, HC(6')); 3.80 (*dd*, ^2J = 11.7, ^3J = 1.9, HC(6')); 3.36 (*ddd*, *d br-like*, ^3J = 9.1, HC(5')); 2.90 (*dd*, ^2J = 14.8, ^3J = 4.8, HC(3')); 2.65 (*dd*, ^2J = 14.8, ^3J = 5.8, HC(3')); 1.09–0.95 (*m*, 28 H (^{i}Pr)).

^{13}C-NMR (100.6 MHz, $CDCl_3$): 197.8 (C(2')); 151.6, 150.1 (C(2), C(4)); 135.9 (C(6)); 101.6 (C(1')); 95.4 (C(5)); 75.3 (C(4')), 67.8 (C(5')); 62.3

(C(6')); 48.6 (C(3')); 17.4, 17.2, 17.2, 17.1 ($CH_3(^iPr)$); 13.3, 13.0, 12.5, 12.4 ($CH(^iPr)$).

ESI-MS: 485.7 (14, $[M+H]^+$); 507.7 (100, $[M+Na]^+$); 523.8 (6, $[M+K]^+$).

6.19 Synthesis of CE PA dU (31)

According to the literature,[48,187] Cyanoethylphosphoramidite dU (**31**) was synthesized, in two steps, from commercially available 2'-Deoxyuridine in good yields.

6.20 Synthesis of Oligonucleotides

According to *Applied Biosystems*, oligonucleotides were synthesized on a DNA-synthesizer (*Expedite* NAS 8905, *Applied Biosystems*) in 0.4 μmol scale using standard specifications and protocol for automated 3'→5' synthesis. Following modifications were made: In order to spend less building blocks, the phosphoramidite concentrations were reduced from 0.1M to 0.05M. Furthermore, the more acidic and more potent 5-(ethylthio)-1*H*-tetrazole was used as activator, instead of 1*H*-tetrazole. In addition, the coupling times for our artificial nucleotide dD were elongated from 90 sec. to 12 min. Oligonucleotide syntheses were performed starting from commercially available nucleosides, bound to *controled pore glas* 500 (CPG 500, *Glen research*). The corresponding commercially available cyanoethyl phosphoramidites were used for the insertion of dA, dG and dC.

The oligonucleotide elongation cycle consisted of four steps:

1) Acidic cleavage of the 5'-end DMT protecting group of the support-bound nucleoside (respectively oligonucleotide).
2) Coupling with 5-(ethylthio)-1*H*-tetrazole activated (2-cyanoethyl)-phosphoramidite.

3) Capping of uncoupled 5'-OH-groups with Ac_2O.
4) Oxidation of the phospite to phosphate with *aq.* I_2-solution.

After finishing chain elongation and cleavage of the 5'-terminal DMT protection group, the oligonucleotides were treated with oversaturated *aq.* NH_3-solution at 50 °C for 20 h to release it from solid support and to chip all nucleobase and phosphate protecting groups. The crude oligonucleotides were purified by precipitation from 70% EtOH/ 30% 0.3M *aq.* Na-acetate, pH = 4.8 and analysed by HPLC and MALDI-TOF-MS (matrix-assisted laser desorption/ionization time-of-flight MS, Matrix: 3-HPA, if not otherwise noted). HPLC was performed on a *DEAE*-ion exchange (IE) column (*Nucleogen-DEAE* 60-7, *Macherey-Nagel*) with a linear KCl-gradient in 20 mM KH_2PO_4, pH = 6.0, in H_2O/CH_3CN 4:1. If necessary oligonucleotides were purified by the same HPLC method. Yields were determined by photometric DMT cation detection in DMT cleavage fractions, collected from synthesizer. In addition, the results were confirmed by determinations of oligonucleotide concentration after work up and purification. The oligonucleotide concentrations were determined by the *Lambert-Beer* law measuring the optical density at 260 nm and assuming the extinction coefficients ε to be the sum of ε of all mononucleosides at the same wavelength. These measurements were carried out at 90 °C in oder to ensure this assumption to be true, because stacking of the nucleobases, and, therefore, hypochromicity are suspended at this temperature.

6.20.1 5'–CGCAUGAGUACGC–3' (A1)

Scale: 0.4 μmol dC-CPG 500

Synthesis protocol: 1 μmol

Yield: 90%

MALDI-TOF-MS: 3930 (77, $[M+11H]^-$); 3952 (100, $[M+Na+10H]^-$); 3972 (82, $[M+2Na+9H]^-$); 3996 (49, $[M+3Na+8H]^-$).

HPLC: t_R = 25.04 min, k' = 20.40.

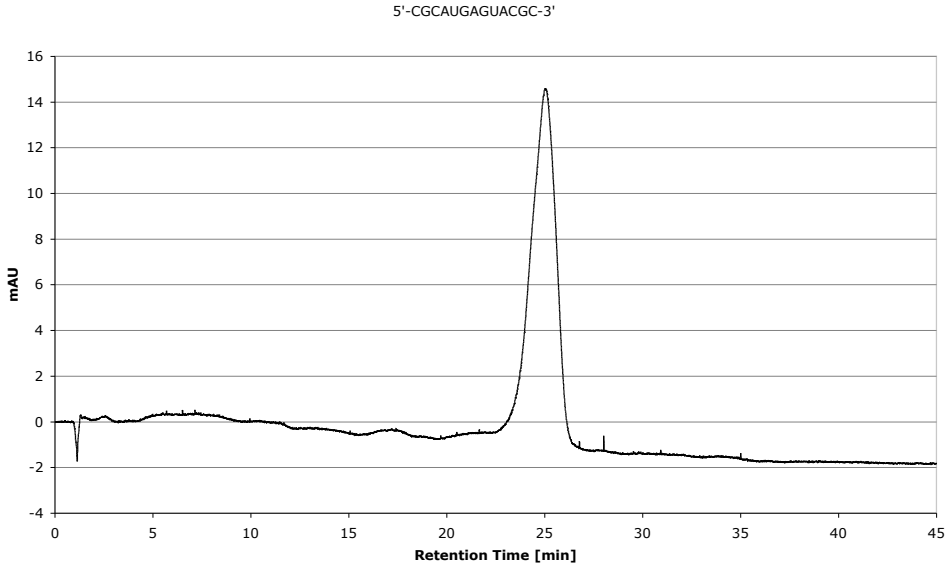


Figure 6.1: HPLC of crude A1.

6.20.2 5'–CGCAUGDGUACGC–3' (A2)

Scale: 0.4 μmol dC-CPG 500

Synthesis protocol: 1 μmol

Yield: 37.3% [3]

dD-CE PA coupling: - Coupling time: 12 min
- Yield: 75.2%

MALDI-TOF-MS: 3905 (91, $[M+11H]^-$); 3927 (100, $[M+Na+10H]^-$); 3949 (65, $[M+2Na+9H]^-$); 3971 (49, $[M+3Na+8H]^-$); 3993 (12, $[M+4Na+7H]^-$).

[3]Low yielded couplings with dD-CE PA and dG-CE PA.

HPLC: $t_R = 25.91$ min, k' = 24.16.

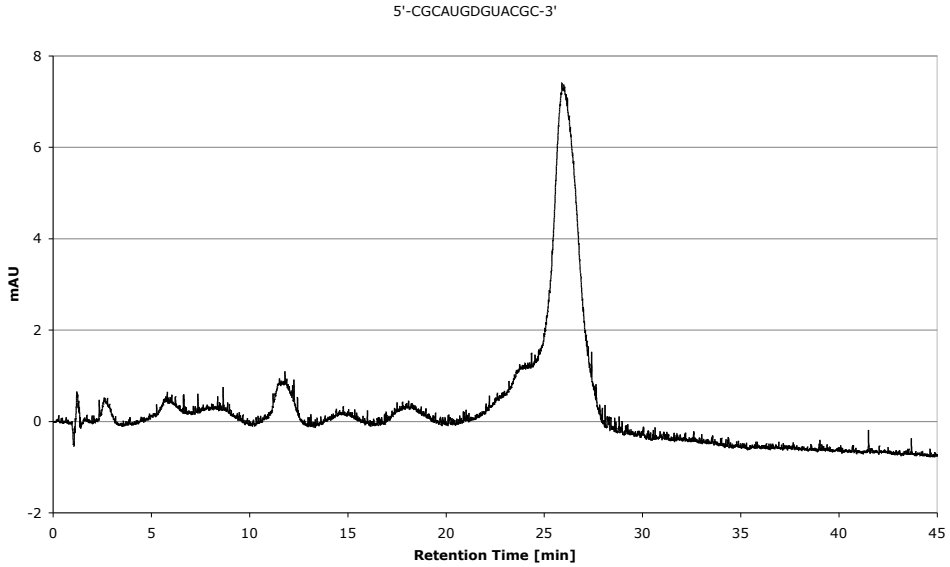


Figure 6.2: HPLC of crude A2.

6.20.3 5'–CGCAUGGGUACGC–3' (A3)

Scale: 0.4 μmol dC-CPG 500

Synthesis protocol: 1 μmol

Yield: 95%

MALDI-TOF-MS: 3946 (63, $[M+11H]^-$); 3968 (100, $[M+Na+10H]^-$); 3990 (78, $[M+2Na+9H]^-$); 4012 (45, $[M+3Na+8H]^-$).

HPLC: $t_R = 24.34$ min, k' = 19.80.

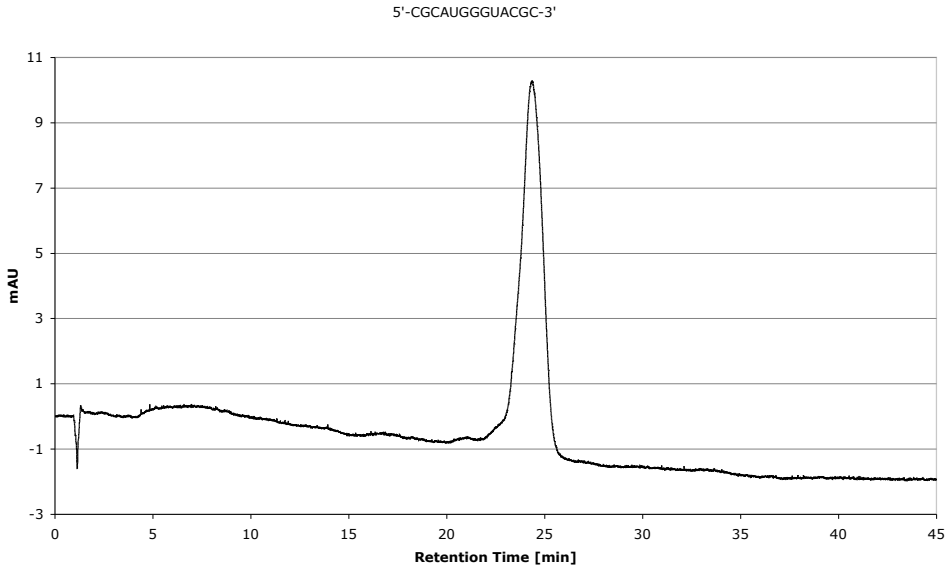


Figure 6.3: HPLC of crude A3.

6.20.4 5'–CGCAUGUGUACGC–3' (A4)

Scale: 0.4 μmol dC-CPG 500

Synthesis protocol: 1 μmol

Yield: 96%

MALDI-TOF-MS: 3907 (100, $[M+11H]^-$); 3929 (8, $[M+Na+10H]^-$).

HPLC: t_R = 24.22 min, k' = 19.70.

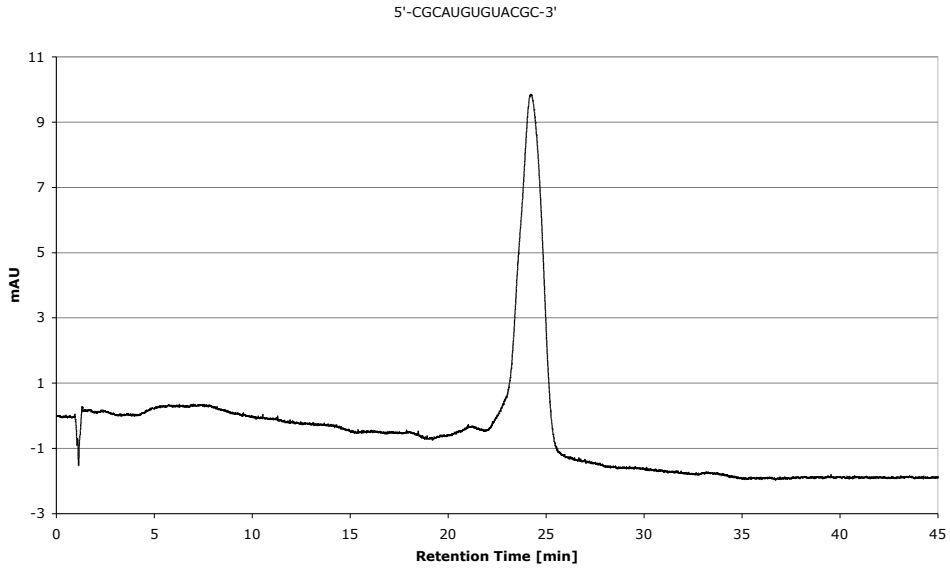


Figure 6.4: HPLC of crude A4.

6.20.5 5'–CGCAUGCCGUACGC–3' (A5)

Scale: 0.4 μmol dC-CPG 500

Synthesis protocol: 1 μmol

Yield: 99%

MALDI-TOF-MS: 3906 (98, $[M+11H]^-$); 3928 (100, $[M+Na+10H]^-$); 3950 (61, $[M+2Na+9H]^-$); 3972 (30, $[M+3Na+8H]^-$); 3994 (13, $[M+4Na+7H]^-$).

HPLC: t_R = 24.47 min, k' = 19.91.

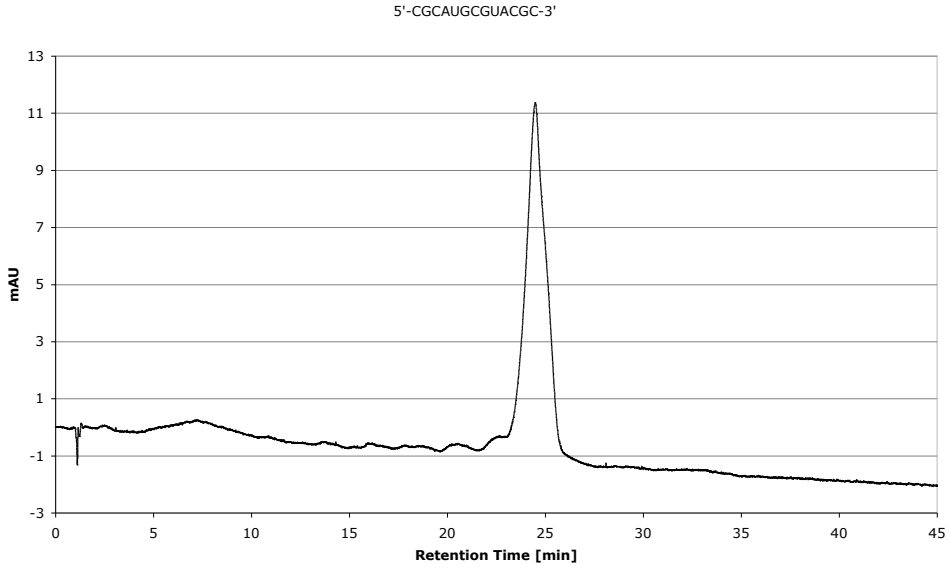


Figure 6.5: HPLC of crude A5.

6.20.6 5'–GCGUACACAUGCG–3' (B1)

Scale: 0.4 μmol dG-CPG 500

Synthesis protocol: 1 μmol

Yield: 90%

MALDI-TOF-MS: 3930 (100, $[M+11H]^-$); 3952 (34, $[M+Na+10H]^-$); 3974 (10, $[M+2Na+9H]^-$).

HPLC: t_R = 25.36 min, k' = 20.68.

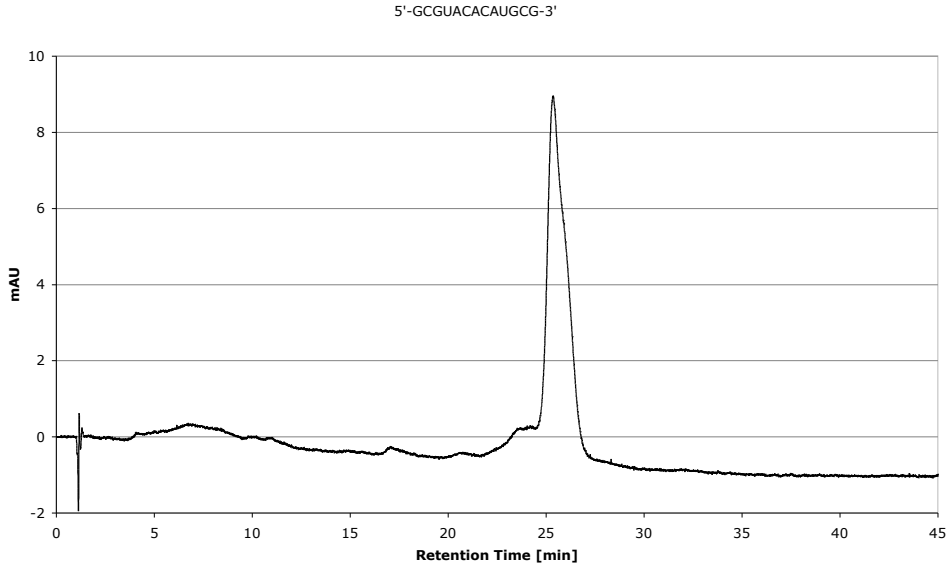


Figure 6.6: HPLC of crude B1.

6.20.7 5'–GCGUACDCAUGCG–3' (B2)

Scale: 0.4 μmol dG-CPG 500

Synthesis protocol: 1 μmol

Yield: 67.1% [4]

dD-CE PA coupling: - Coupling time: 12 min
- Yield: 77.2%

MALDI-TOF-MS: 3904 (100, $[M+11H]^-$); 3926 (56, $[M+Na+10H]^-$); 3948 (16, $[M+2Na+9H]^-$); 3970 (5, $[M+3Na+8H]^-$).

HPLC: t_R = 25.97 min, k' = 22.40.

[4]Low yielded couplings with dD-CE PA and dG-CE PA.

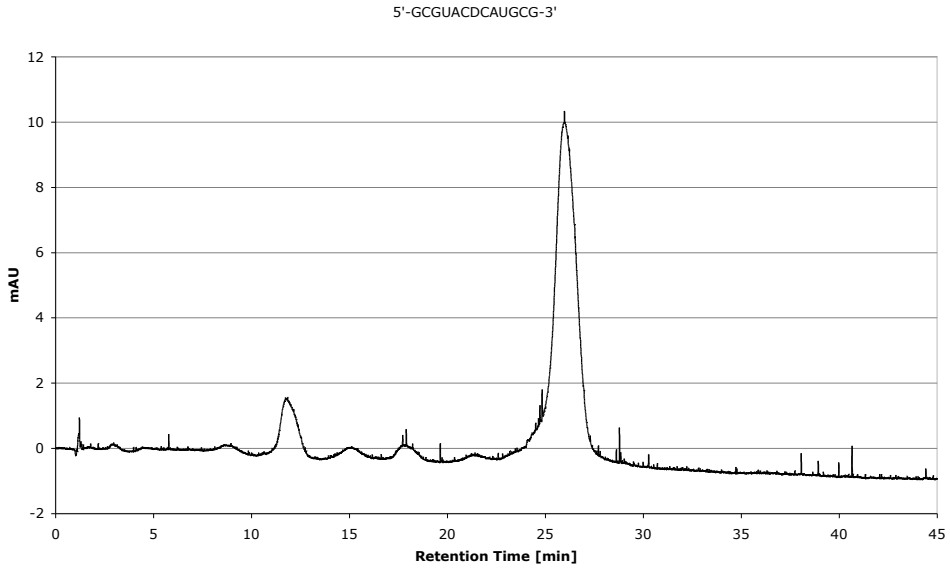


Figure 6.7: HPLC of crude B2.

6.20.8 5'–GCGUACGCAUGCG–3' (B3)

Scale: 0.4 μmol dG-CPG 500

Synthesis protocol: 1 μmol

Yield: 78%

MALDI-TOF-MS: 3946 (100, $[M+11H]^-$); 3968 (45, $[M+Na+10H]^-$); 3990 (13, $[M+2Na+9H]^-$); 4012 (7, $[M+3Na+8H]^-$).

HPLC: t_R = 24.41 min, k' = 19.86.

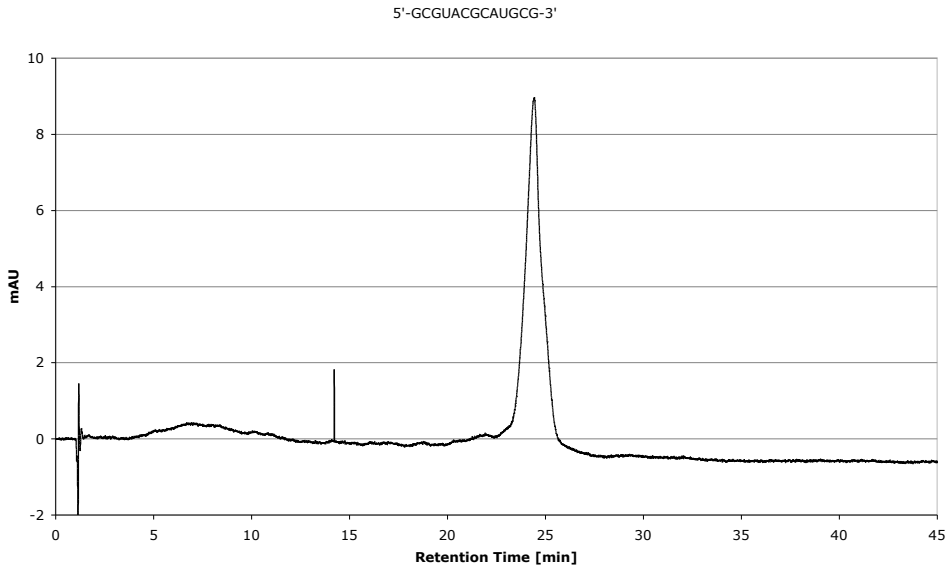


Figure 6.8: HPLC of crude B3.

6.20.9 5'–GCGUACUCAUGCG–3' (B4)

Scale: 0.4 μmol dG-CPG 500

Synthesis protocol: 1 μmol

Yield: 94%

MALDI-TOF-MS: 3907 (100, $[M+11H]^-$); 3929 (18, $[M+Na+10H]^-$).

HPLC: t_R = 23.92 min, k' = 19.44.

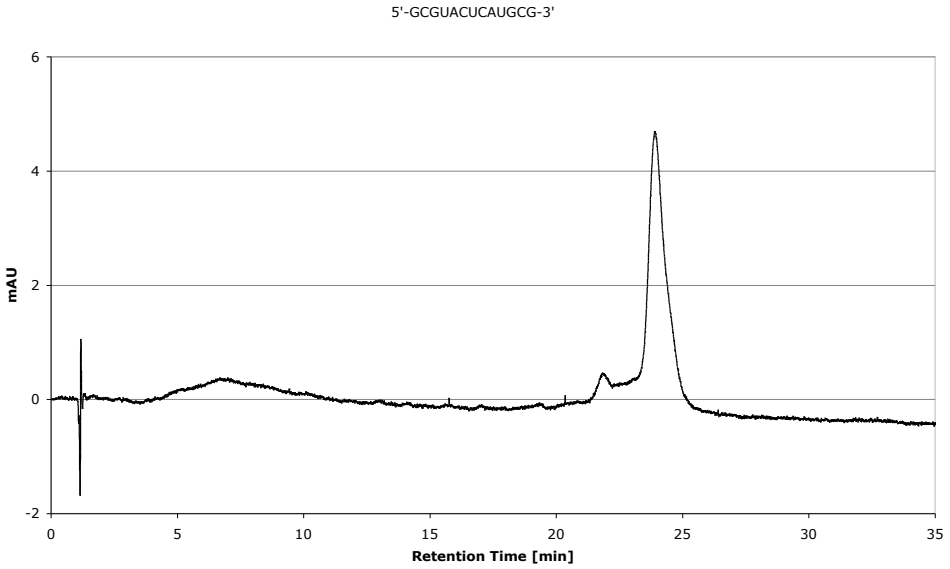


Figure 6.9: HPLC of crude B4.

6.20.10 5'–GCGUACCCAUGCG–3' (B5)

Scale: 0.4 μmol dG-CPG 500

Synthesis protocol: 1 μmol

Yield: 99%

MALDI-TOF-MS: 3906 (100, $[M+11H]^-$); 3928 (37, $[M+Na+10H]^-$).

HPLC: t_R = 25.45 min, k' = 20.75.

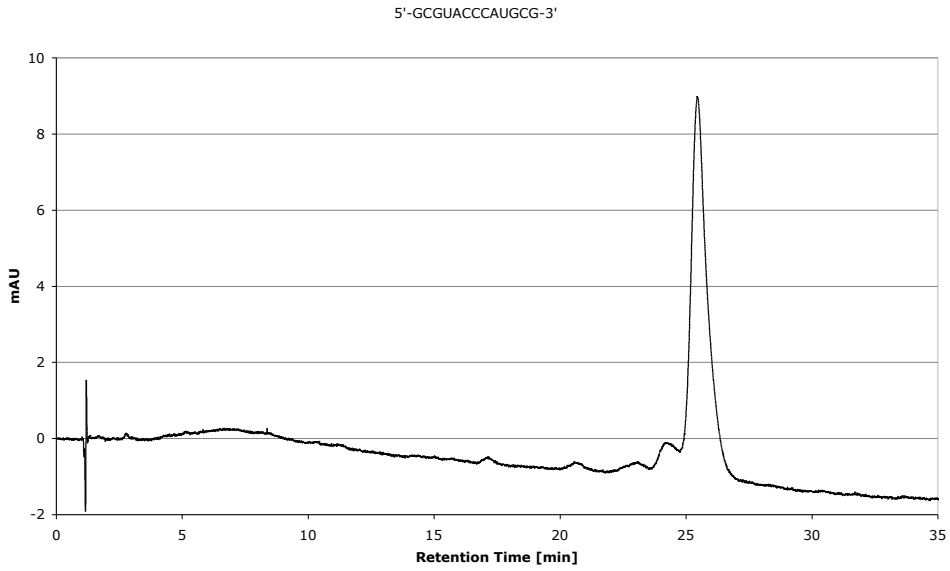


Figure 6.10: HPLC of crude B5.

6.20.11 5'–GUACGC–3' (32)

Scale: 0.4 μmol dC-CPG 500

Synthesis protocol: 1 μmol

Yield: 53% [5]

MALDI-TOF-MS[6]: 1777 (100, $[M+4H]^-$); 1799 (5, $[M+Na+3H]^-$); 1815 (2, $[M+K+3H]^-$).

HPLC: t_R = 12.16 min, k' = 11.41.

[5]Low yielded couplings with dG-CE PA.

[6]Matrix: ATT.

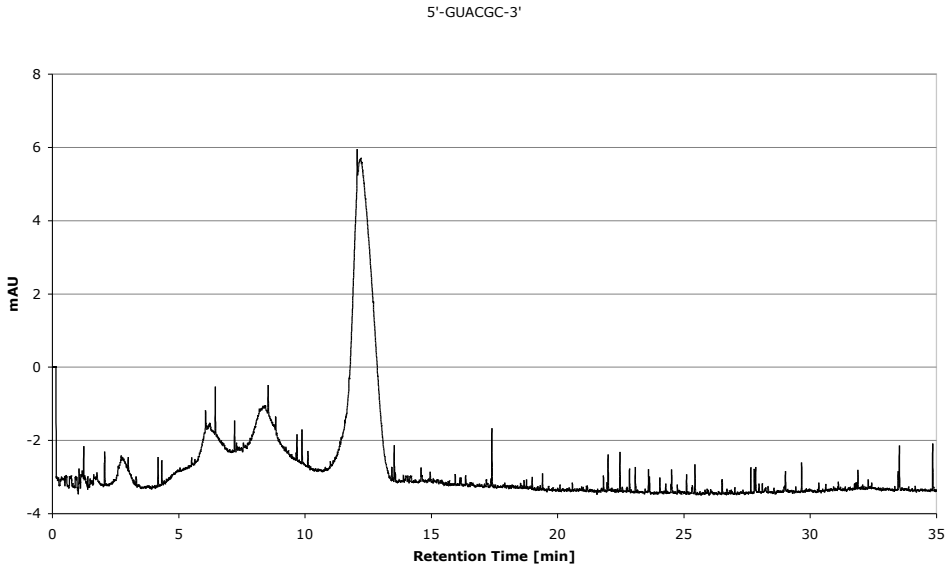


Figure 6.11: HPLC of crude **32**.

6.20.12 5'–CAUGCG–3' (33)

Scale: 0.4 μmol dG-CPG 500

Synthesis protocol: 1 μmol

Yield: 81% [7]

MALDI-TOF-MS[8]: 1777 (100, $[M+4H]^-$); 1799 (10, $[M+Na+3H]^-$); 1815 (2, $[M+K+3H]^-$).

HPLC: t_R = 12.33 min, k' = 11.58.

[7]Low yielded couplings with dG-CE PA.

[8]Matrix: ATT.

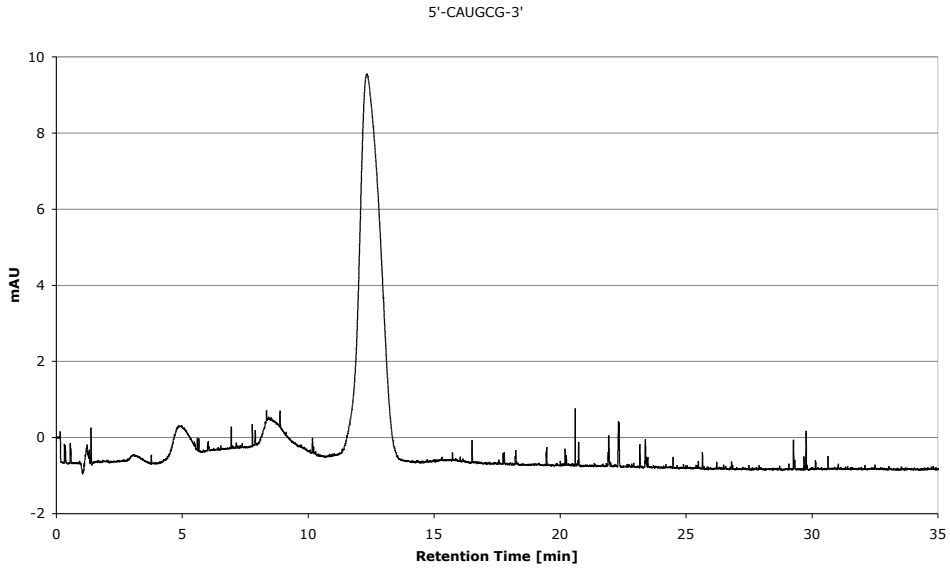


Figure 6.12: HPLC of crude **33**.

Chapter 7

UV–Melting Curves

7.1 Base pairing of A1B1

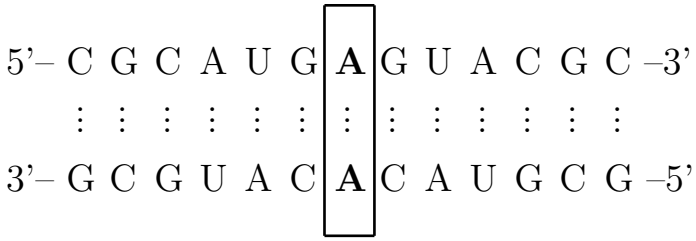


Figure 7.1: Double stranded oligonucleotide A1B1.

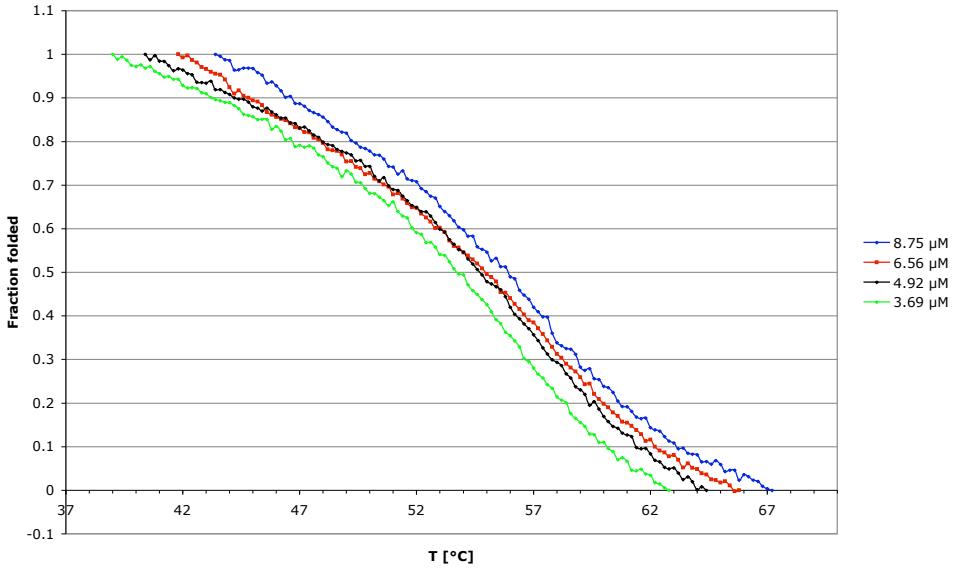


Figure 7.2: Fraction folded plot of A1B1.

Table 7.1: T_m and thermodynamic data of A1B1.

Concentration [μM]	8.75	6.56	4.92	3.69
T_m [°C]	56.2	55.4	54.6	53.8

	$\Delta G°$[kcal/mol]	$\Delta H°$[kcal/mol]	$\Delta S°$[cal/mol]	R^2
ln C_{Total} vs. $1/T_m$	-14.9	-76.4	-206	0.999
$1/T$ vs. ln K_a	-14.9 ± 0.1	-76.8 ± 1.5	-208 ± 5	

7.2 Base pairing of A1B2

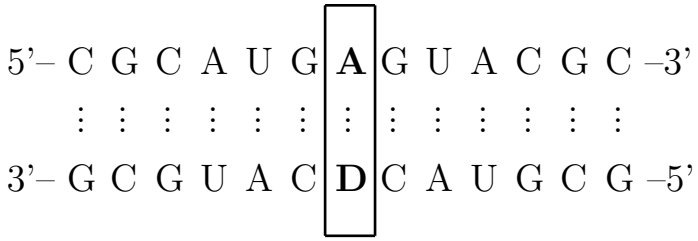


Figure 7.3: Double stranded oligonucleotide A1B2.

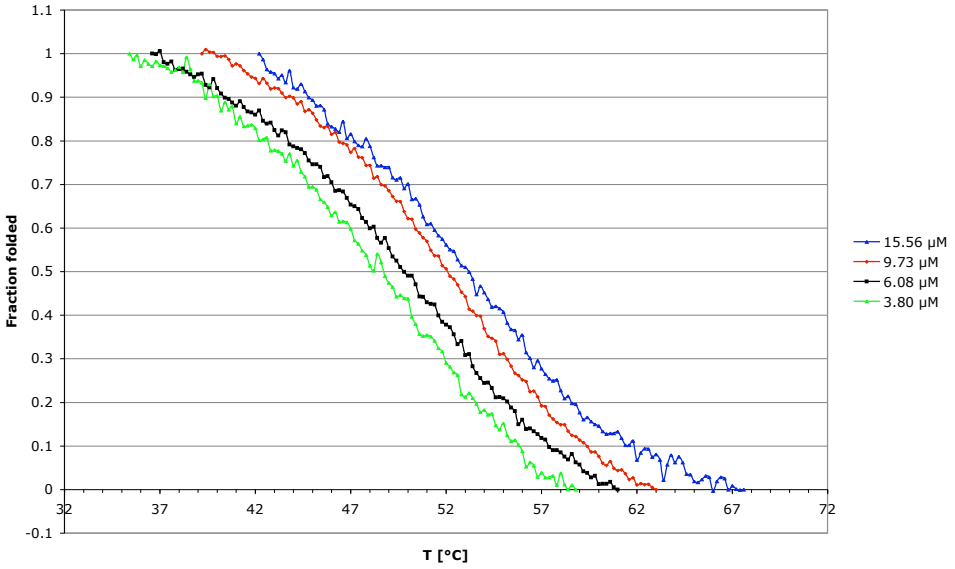


Figure 7.4: Fraction folded plot of A1B2.

Table 7.2: T_m and thermodynamic data of A1B2.

Concentration [μM]	15.56	9.75	6.08	3.80
T_m [°C]	52.5	51.6	50.0	48.6

	$\Delta G°$ [kcal/mol]	$\Delta H°$ [kcal/mol]	$\Delta S°$ [cal/mol]	R^2
ln C_{Total} vs. $^1/T_m$	-13.6	-73.5	-201	0.988
$^1/T$ vs. ln K_a	-13.9 ± 0.2	-76.8 ± 1.0	-211 ± 3	

7.3 Base pairing of A1B3

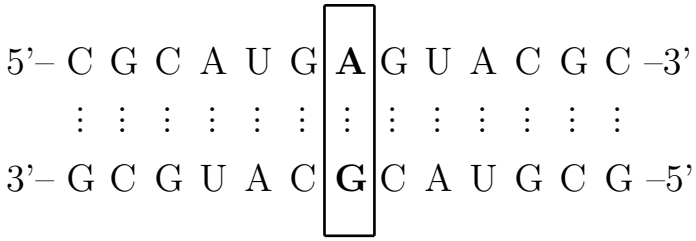


Figure 7.5: Double stranded oligonucleotide A1B3.

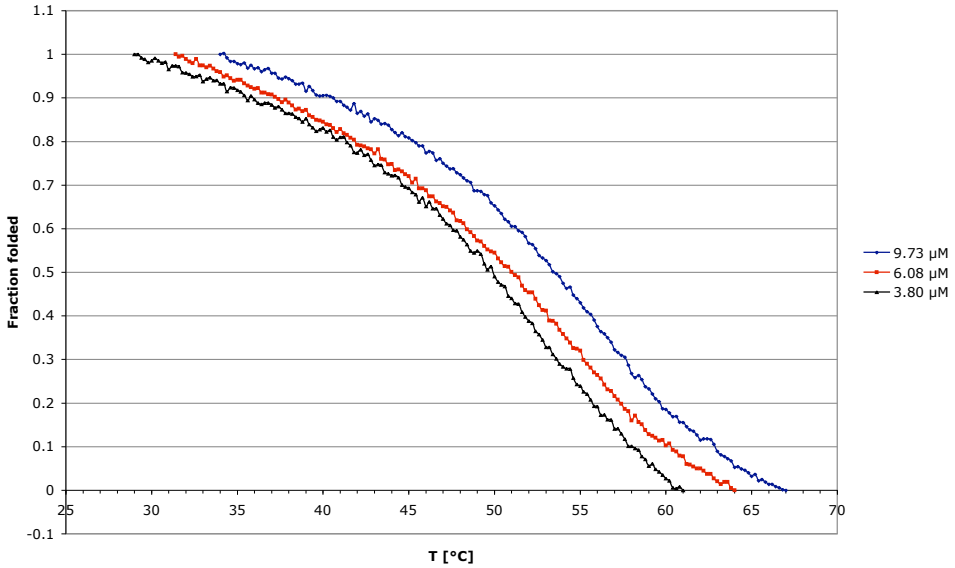


Figure 7.6: Fraction folded plot of A1B3.

Table 7.3: T_m and thermodynamic data of A1B3.

Concentration [μM]	9.73	6.08	3.8
T_m [°C]	53.4	51.4	49.8

	$\Delta G°$[kcal/mol]	$\Delta H°$[kcal/mol]	$\Delta S°$[cal/mol]	R^2
ln C_{Total} vs. $1/T_m$	-12.5	-55.2	-143	0.996
$1/T$ vs. ln K_a	-12.6 ± 0.3	-57.5 ± 2.50	-151 ± 7	

7.4 Base pairing of A1B4

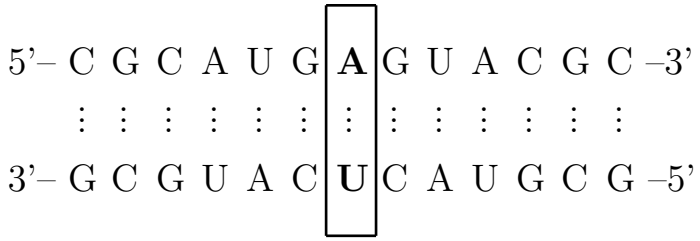


Figure 7.7: Double stranded oligonucleotide A1B4.

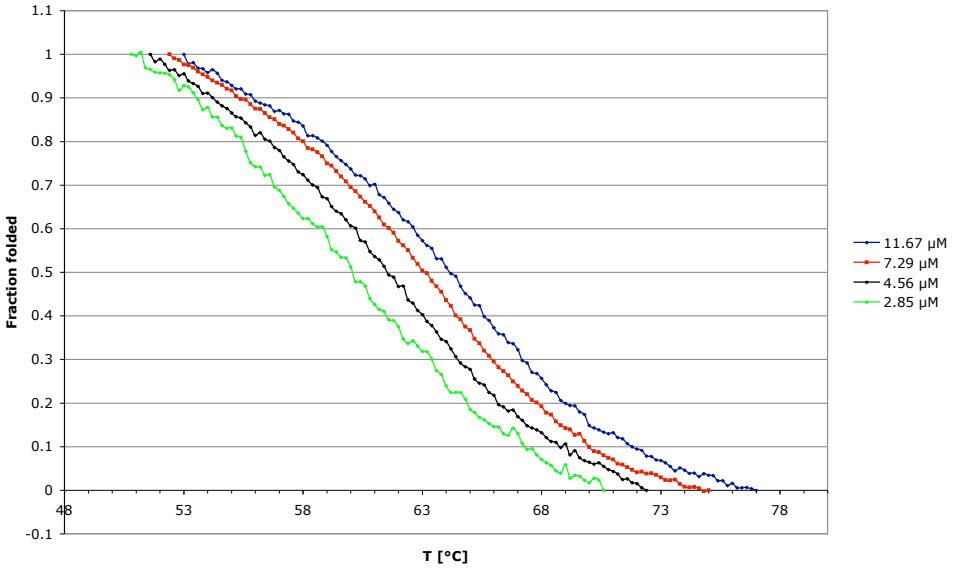


Figure 7.8: Fraction folded plot of A1B4.

Table 7.4: T_m and thermodynamic data of A1B4.

Concentration [μM]	11.67	7.29	4.56	2.85
T_m [°C]	64.1	63	61.8	60.8

	$\Delta G°$[kcal/mol]	$\Delta H°$[kcal/mol]	$\Delta S°$[cal/mol]	R^2
ln C_{Total} vs. $1/T_m$	-18.5	-94.5	-255	0.999
$1/T$ vs. ln K_a	-18.5 ± 0.2	-94.7 ± 1.1	-256 ± 3	

7.5 Base pairing of A1B5

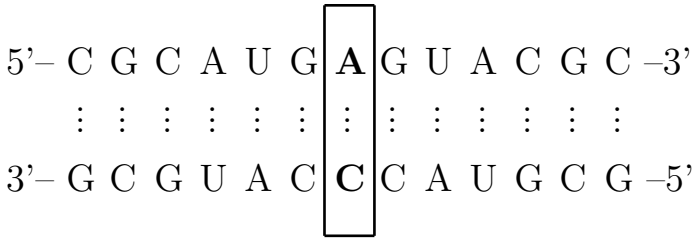


Figure 7.9: Double stranded oligonucleotide A1B5.

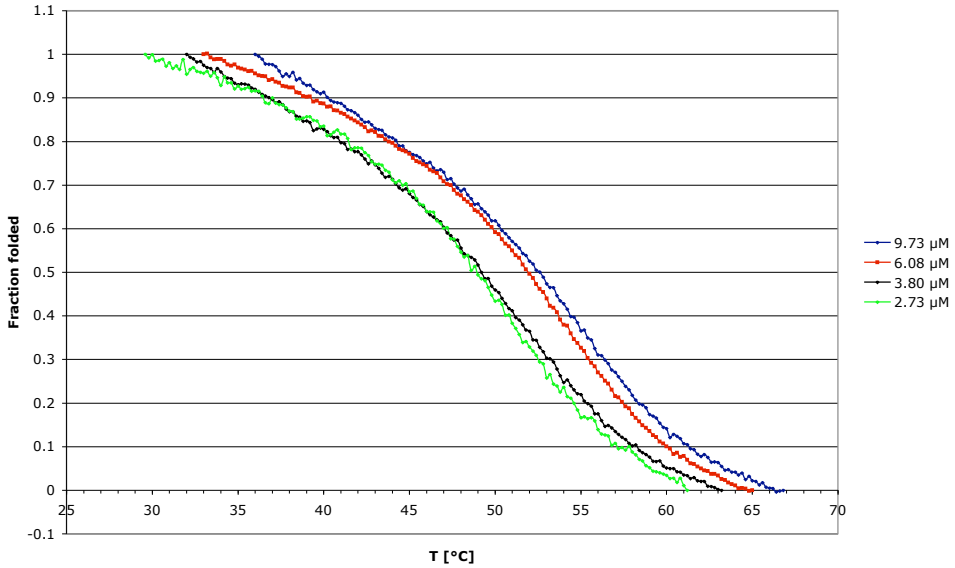


Figure 7.10: Fraction folded plot of A1B5.

Table 7.5: T_m and thermodynamic data of A1B5.

Concentration [μM]	9.73	6.08	3.80	2.73
T_m [°C]	52.8	51.6	49.6	48.6

	$\Delta G°[kcal/mol]$	$\Delta H°[kcal/mol]$	$\Delta S°[cal/mol]$	R^2
ln C_{Total} vs. $1/T_m$	-13.1	-62.0	-164	0.990
$1/T$ vs. ln K_a	-13.0 ± 0.3	-62.7 ± 1.9	-167 ± 6	

7.6 Base pairing of A2B1

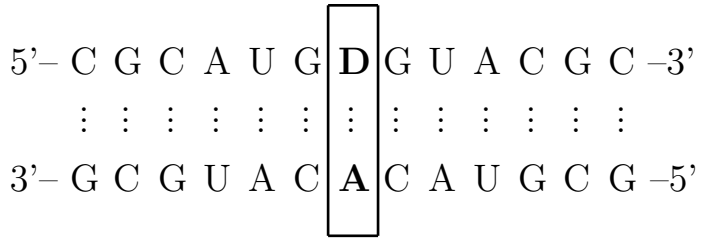


Figure 7.11: Double stranded oligonucleotide A2B1.

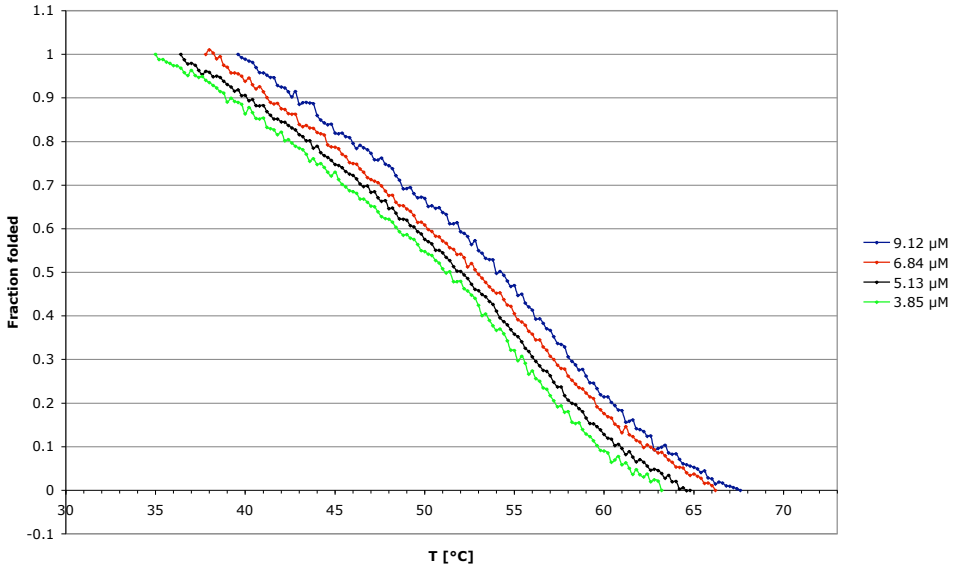


Figure 7.12: Fraction folded plot of A2B1.

Table 7.6: T_m and thermodynamic data of A2B1.

Concentration [μM]	9.12	6.84	5.13	3.85
T_m [°C]	54.2	52.9	52.0	50.9

	$\Delta G°$[kcal/mol]	$\Delta H°$[kcal/mol]	$\Delta S°$[cal/mol]	R^2
ln C_{Total} vs. $1/T_m$	-12.8	-56.7	-147	0.996
$1/T$ vs. ln K_a	-12.7 ± 0.1	-57.0 ± 0.8	-148 ± 3	

7.7 Base pairing of A2B2

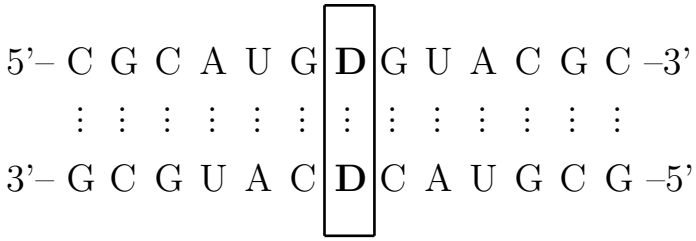


Figure 7.13: Double stranded oligonucleotide A2B2.

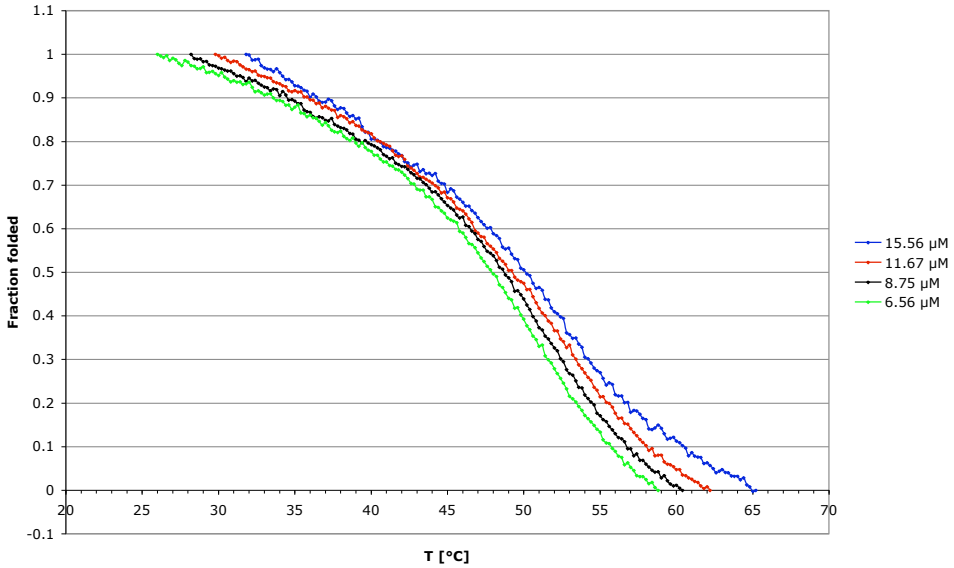


Figure 7.14: Fraction folded plot of A2B2.

Table 7.7: T_m and thermodynamic data of A2B2.

Concentration [μM]	15.56	11.67	8.75	6.56
T_m [°C]	50.8	49.6	48.8	47.6

	$\Delta G°[kcal/mol]$	$\Delta H°[kcal/mol]$	$\Delta S°[cal/mol]$	R^2
ln C_{Total} vs. $1/T_m$	-11.9	-56.7	-151	0.994
$1/T$ vs. ln K_a	-11.8 ± 0.1	-56.1 ± 0.3	-149 ± 1	

7.8 Base pairing of A2B3

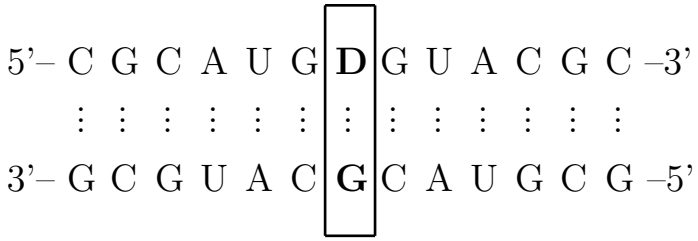


Figure 7.15: Double stranded oligonucleotide A2B3.

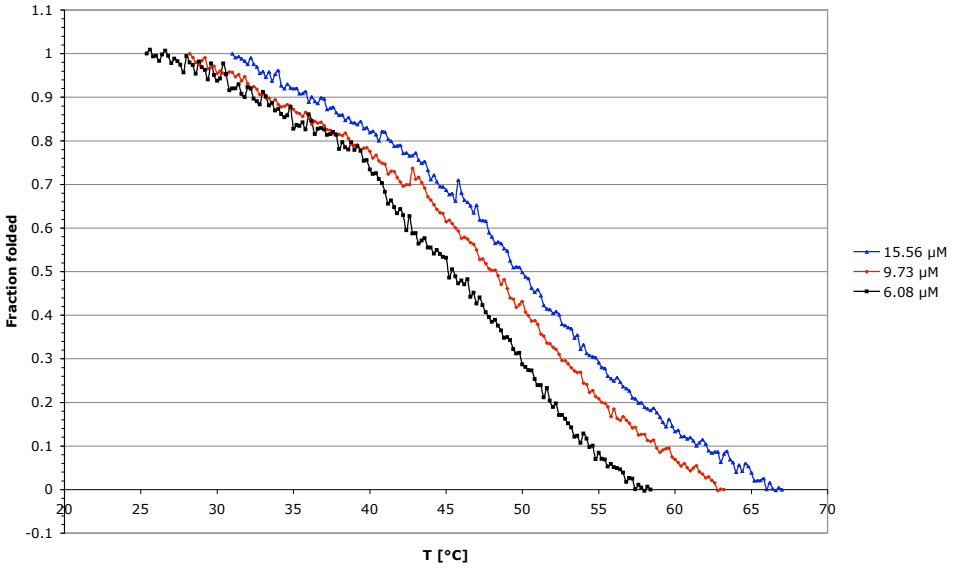


Figure 7.16: Fraction folded plot of A2B3.

Table 7.8: T_m and thermodynamic data of A2B3.

Concentration [μM]	15.56	9.73	6.08
T_m [°C]	49.8	48.0	46.0

	$\Delta G°$[kcal/mol]	$\Delta H°$[kcal/mol]	$\Delta S°$[cal/mol]	R^2
ln C_{Total} vs. $1/T_m$	-11.3	-50.9	-133	0.998
$1/T$ vs. ln K_a	-11.4 ± 0.2	-54.2 ± 3.4	-144 ± 10	

7.9 Base pairing of A2B4

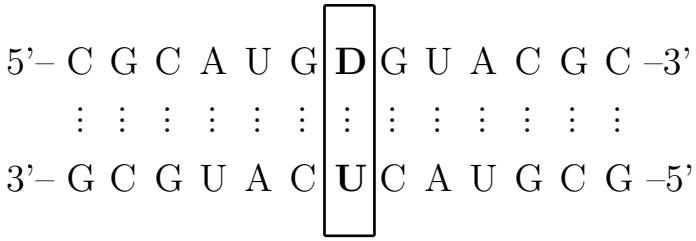


Figure 7.17: Double stranded oligonucleotide A2B4.

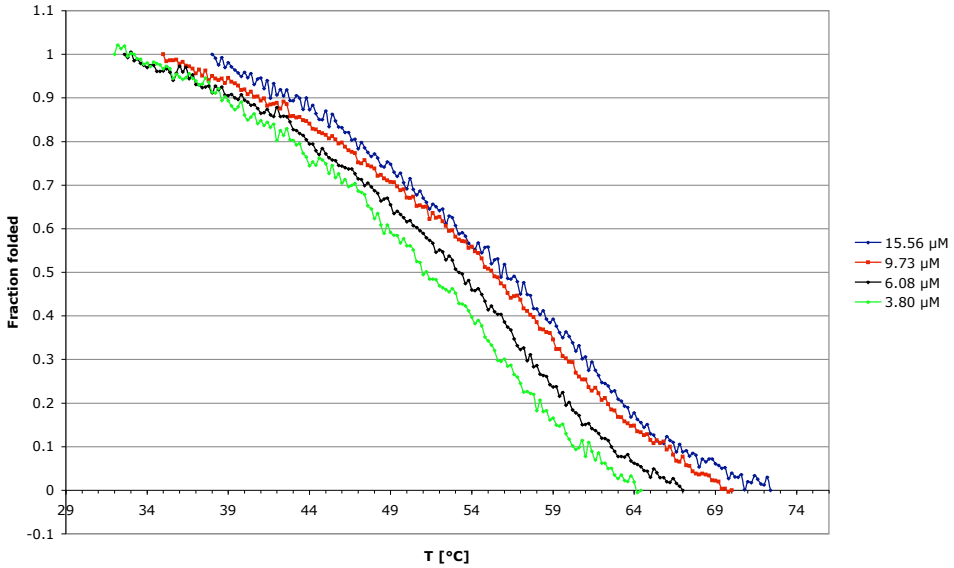


Figure 7.18: Fraction folded plot of A2B4.

Table 7.9: T_m and thermodynamic data of A2B4.

Concentration [μM]	15.56	9.73	6.08	3.80
T_m [°C]	56.6	54.8	53.2	51.3

	$\Delta G°$[kcal/mol]	$\Delta H°$[kcal/mol]	$\Delta S°$[cal/mol]	R^2
ln C_{Total} vs. $1/T_m$	-12.8	-56.7	-147	0.999
$1/T$ vs. ln K_a	-12.6 ± 0.2	-54.3 ± 2.1	-140 ± 6	

7.10 Base pairing of A2B5

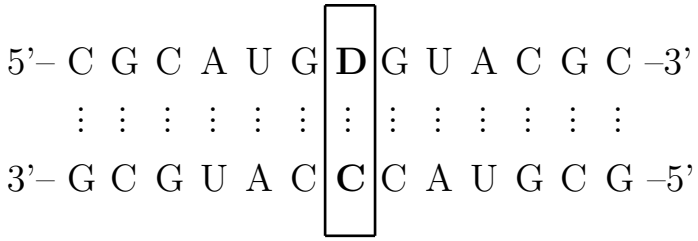


Figure 7.19: Double stranded oligonucleotide A2B5.

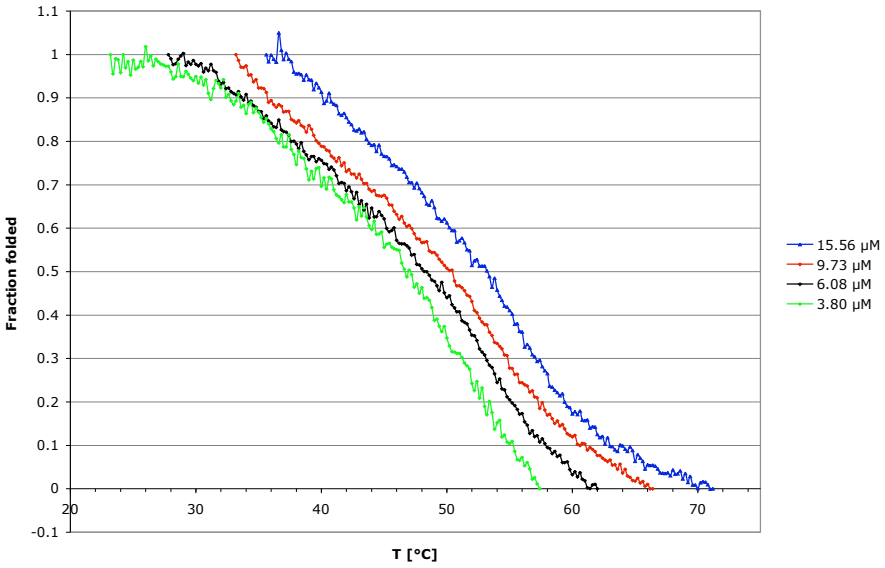


Figure 7.20: Fraction folded plot of A2B5.

Table 7.10: T_m and thermodynamic data of A2B5.

Concentration [μM]	15.56	9.73	6.08	3.80
T_m [°C]	53.0	50.2	48.6	46.6

	$\Delta G°$[kcal/mol]	$\Delta H°$[kcal/mol]	$\Delta S°$[cal/mol]	R^2
ln C_{Total} vs. $1/T_m$	-11.5	-47.3	-120	0.988
$1/T$ vs. ln K_a	-11.4 ± 0.1	-48.3 ± 1.3	-124 ± 4	

7.11 Base pairing of A3B1

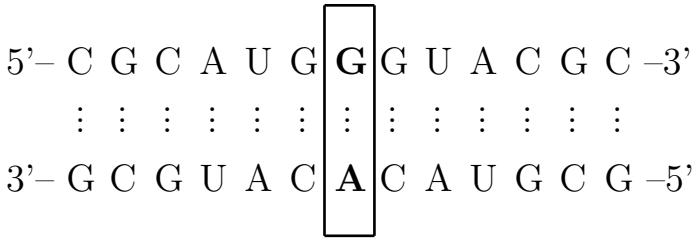


Figure 7.21: Double stranded oligonucleotide A3B1.

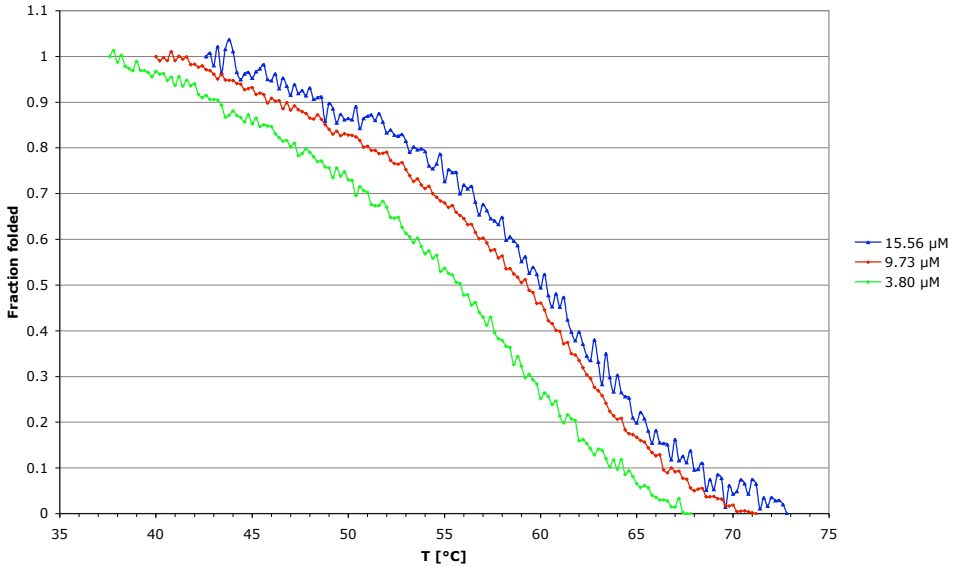


Figure 7.22: Fraction folded plot of A3B1.

Table 7.11: T_m and thermodynamic data of A3B1.

Concentration [μM]	15.56	9.73	3.80
T_m [°C]	60.6	59.0	55.8

	$\Delta G°$[kcal/mol]	$\Delta H°$[kcal/mol]	$\Delta S°$[cal/mol]	R^2
ln C_{Total} vs. $1/T_m$	-14.1	-64.0	-167	0.997
$1/T$ vs. ln K_a	-14.2 ± 0.3	-64.9 ± 2.4	-170 ± 7	

7.12 Base pairing of A3B2

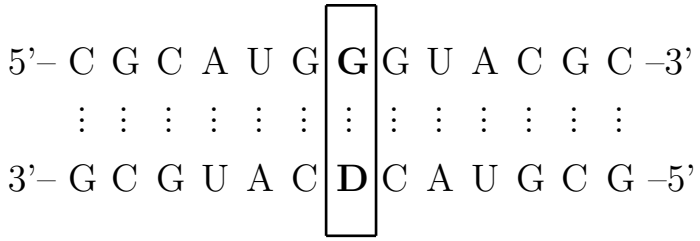


Figure 7.23: Double stranded oligonucleotide A3B2.

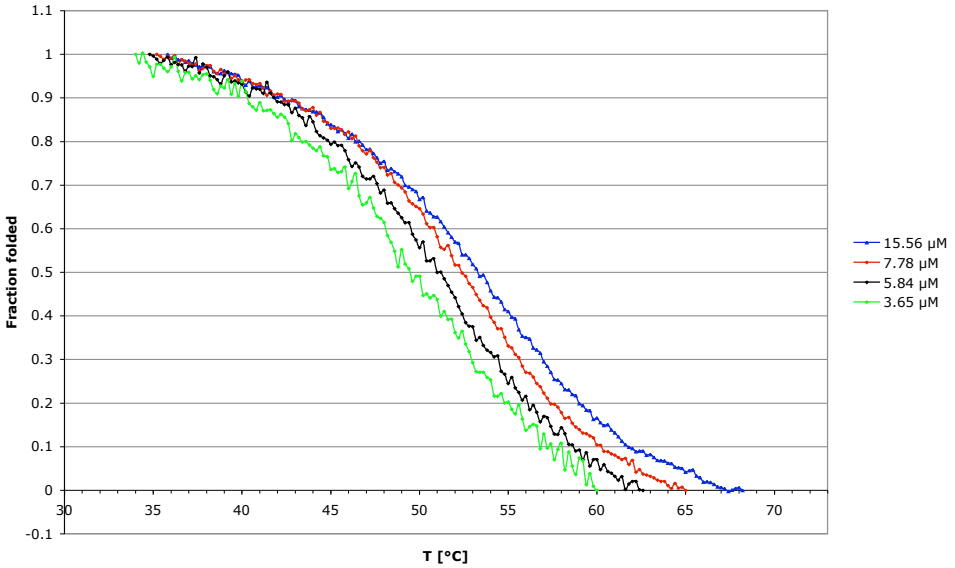


Figure 7.24: Fraction folded plot of A3B2.

Table 7.12: T_m and thermodynamic data of A3B2.

Concentration [μM]	15.56	9.73	6.08	3.80
T_m [°C]	53.8	52.2	51.2	50.0

	$\Delta G°$[kcal/mol]	$\Delta H°$[kcal/mol]	$\Delta S°$[cal/mol]	R^2
ln C_{Total} vs. $1/T_m$	-14.4	-79.4	-218	0.995
$1/T$ vs. ln K_a	-14.2 ± 0.2	-77.1 ± 1.2	-211 ± 4	

7.13 Base pairing of A3B3

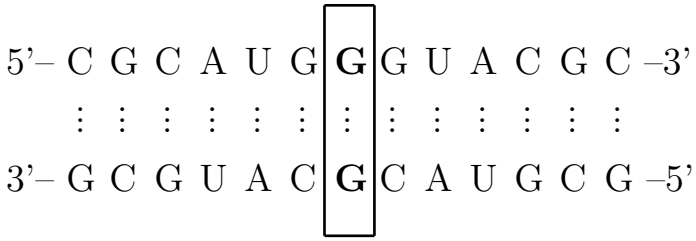


Figure 7.25: Double stranded oligonucleotide A3B3.

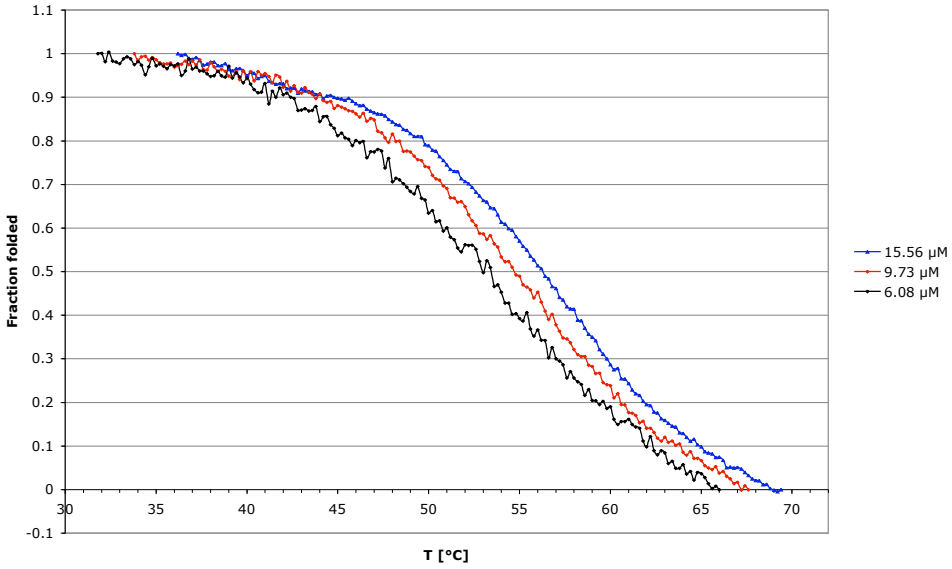


Figure 7.26: Fraction folded plot of A3B3.

Table 7.13: T_m and thermodynamic data of A3B3.

Concentration [μM]	15.56	9.73	6.08
T_m [°C]	56.0	54.6	53.1

	$\Delta G°$[kcal/mol]	$\Delta H°$[kcal/mol]	$\Delta S°$[cal/mol]	R^2
ln C_{Total} vs. $1/T_m$	-13.8	-68.5	-183	0.999
$1/T$ vs. ln K_a	-13.8 ± 0.3	-67.6 ± 2.9	-181 ± 9	

7.14 Base pairing of A3B4

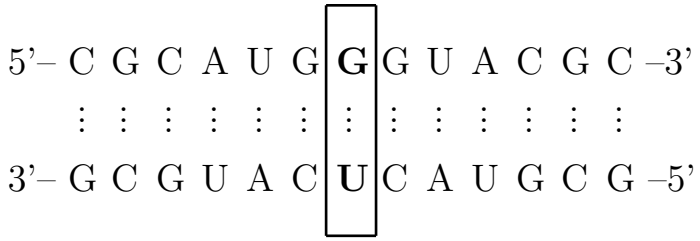


Figure 7.27: Double stranded oligonucleotide A3B4.

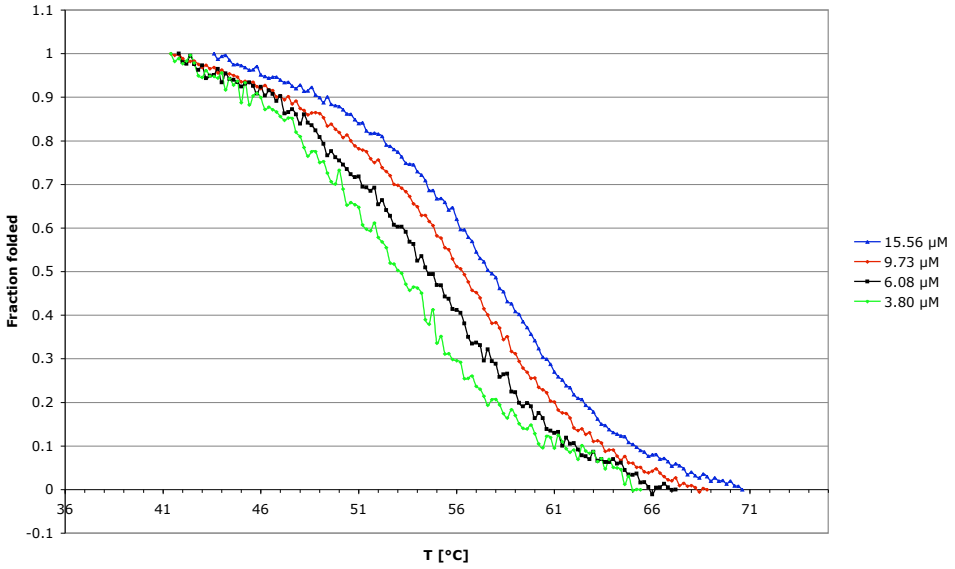


Figure 7.28: Fraction folded plot of A3B4.

Table 7.14: T_m and thermodynamic data of A3B4.

Concentration [μM]	15.56	9.73	6.08	3.80
T_m [°C]	57.4	56.4	55.0	54.0

	$\Delta G°$ [kcal/mol]	$\Delta H°$ [kcal/mol]	$\Delta S°$ [cal/mol]	R^2
ln C_{Total} vs. $1/T_m$	-15.8	-86.3	-237	0.995
$1/T$ vs. ln K_a	-16.0 ± 0.3	-87.9 ± 1.6	-241 ± 5	

7.15 Base pairing of A3B5

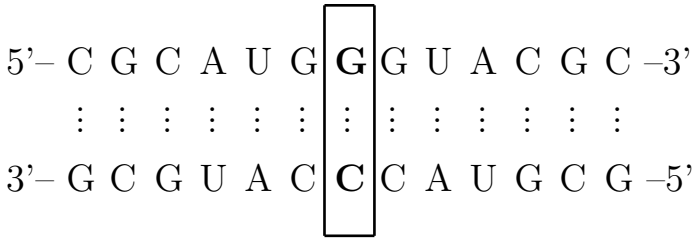


Figure 7.29: Double stranded oligonucleotide A3B5.

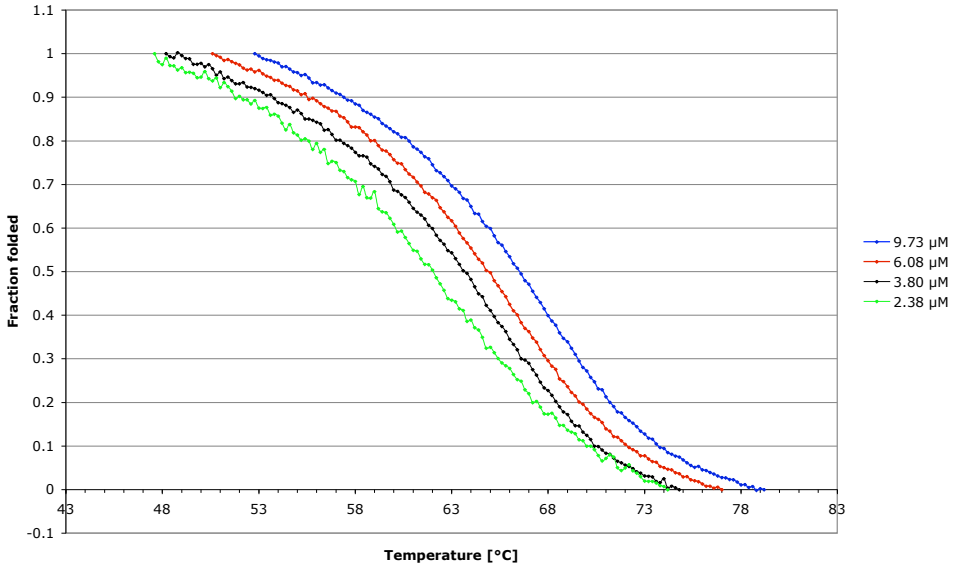


Figure 7.30: Fraction folded plot of A3B5.

Table 7.15: T_m and thermodynamic data of A3B5.

Concentration [μM]	9.73	6.08	3.80	2.38
T_m [°C]	67.1	66.4	64.0	62.4

	$\Delta G°$[kcal/mol]	$\Delta H°$[kcal/mol]	$\Delta S°$[cal/mol]	R^2
ln C_{Total} vs. $1/T_m$	-17.8	-82.7	-218	0.993
$1/T$ vs. ln K_a	-18.2 ± 0.4	-86.3 ± 3.0	-229 ± 8	

7.16 Base pairing of A4B1

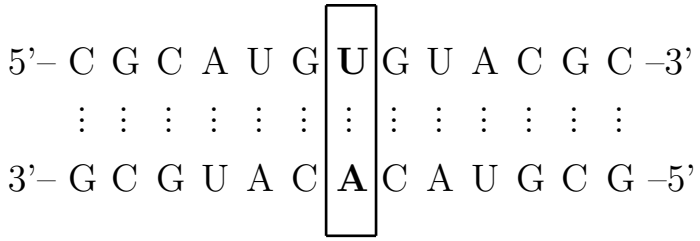


Figure 7.31: Double stranded oligonucleotide A4B1.

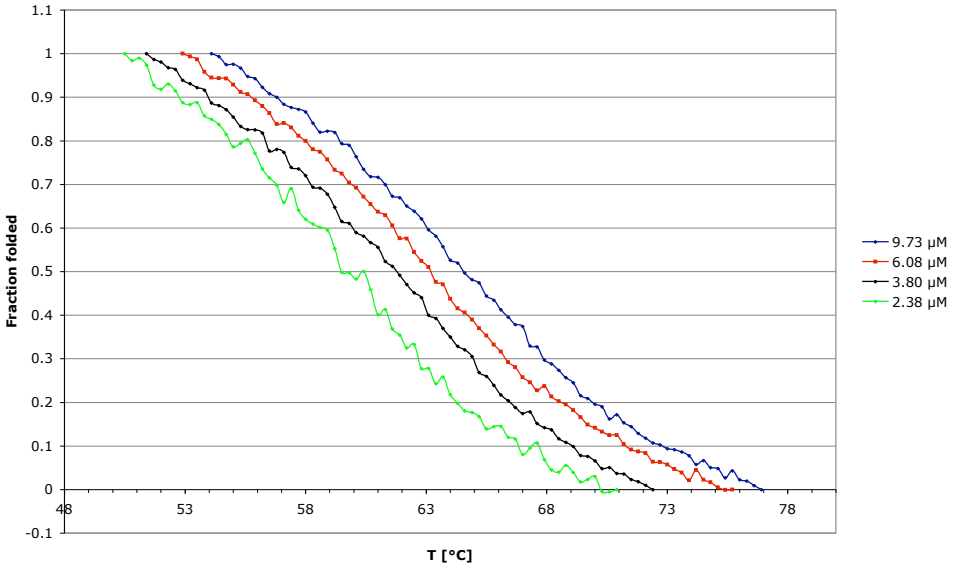


Figure 7.32: Fraction folded plot of A4B1.

Table 7.16: T_m and thermodynamic data of A4B1.

Concentration [μM]	9.73	6.08	3.80	2.38
T_m [°C]	64.5	63.4	61.9	60.9

	$\Delta G°$ [kcal/mol]	$\Delta H°$ [kcal/mol]	$\Delta S°$ [cal/mol]	R^2
ln C_{Total} vs. $1/T_m$	-17.8	-86.3	-230	0.994
$1/T$ vs. ln K_a	-18.1 ± 0.1	-88.6 ± 0.9	-237 ± 3	

7.17 Base pairing of A4B2

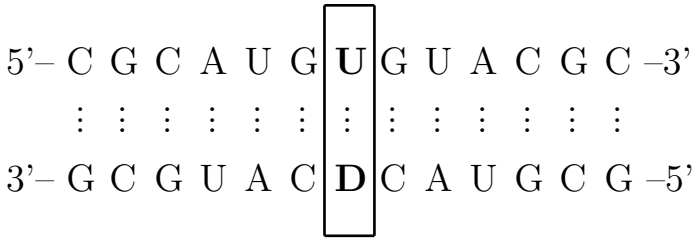


Figure 7.33: Double stranded oligonucleotide A4B2.

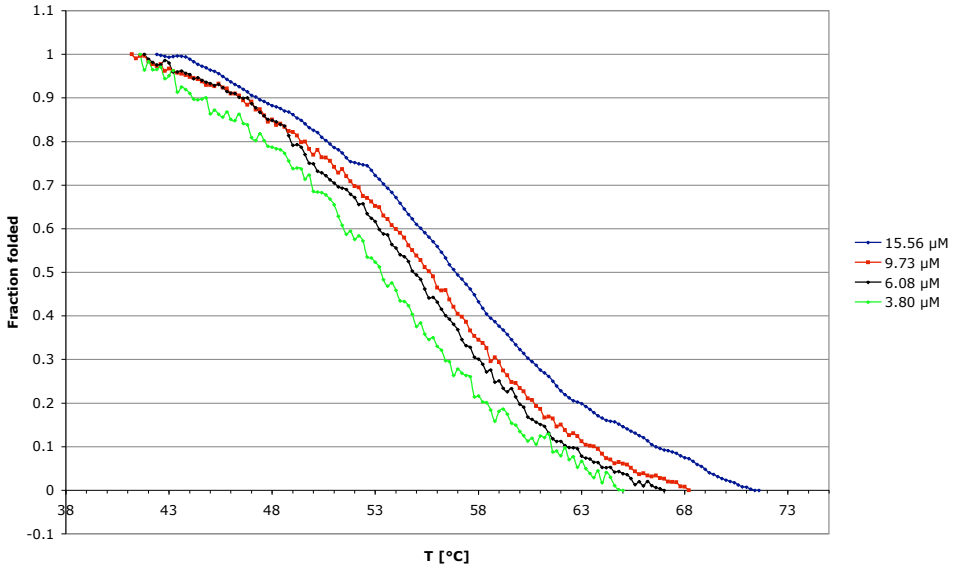


Figure 7.34: Fraction folded plot of A4B2.

Table 7.17: T_m and thermodynamic data of A4B2.

Concentration [μM]	15.56	9.73	6.08	3.80
T_m [°C]	57.0	55.8	54.7	53.3

	$\Delta G°$[kcal/mol]	$\Delta H°$[kcal/mol]	$\Delta S°$[cal/mol]	R^2
ln C_{Total} vs. $1/T_m$	-15.4	-82.7	-226	0.997
$1/T$ vs. ln K_a	-15.2 ± 0.1	-80.9 ± 1.0	-220 ± 3	

7.18 Base pairing of A4B3

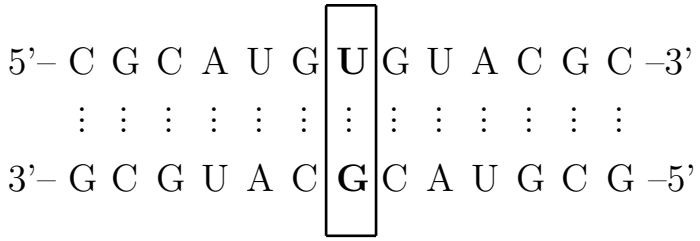


Figure 7.35: Double stranded oligonucleotide A4B3.

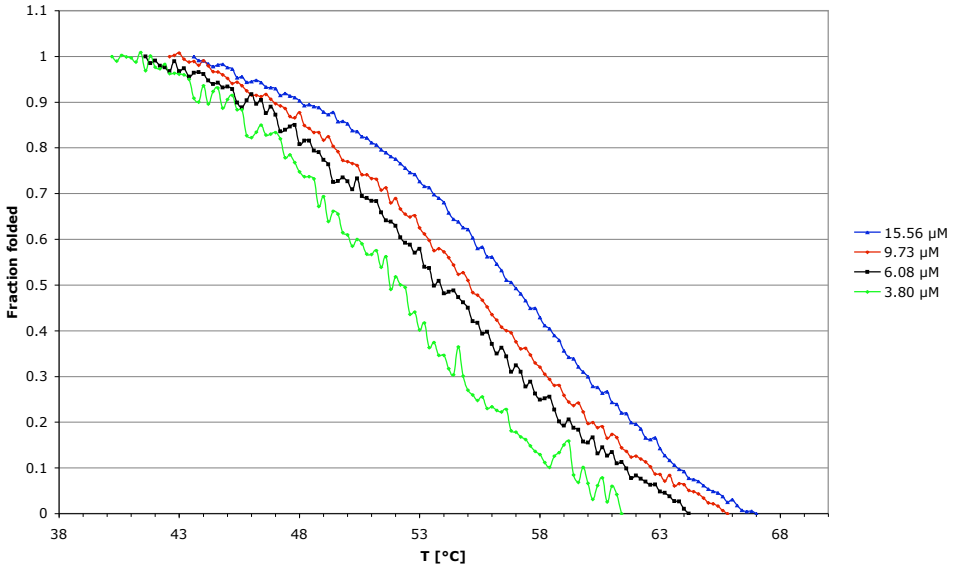


Figure 7.36: Fraction folded plot of A4B3.

Table 7.18: T_m and thermodynamic data of A4B3.

Concentration [μM]	15.56	9.73	6.08	3.80
T_m [°C]	56.2	55.0	53.6	52.6

	$\Delta G°$[kcal/mol]	$\Delta H°$[kcal/mol]	$\Delta S°$[cal/mol]	R^2
ln C_{Total} vs. $1/T_m$	-15.3	-82.7	-226	0.996
$1/T$ vs. ln K_a	-15.5 ± 0.2	-84.1 ± 1.3	-230 ± 4	

7.19 Base pairing of A4B4

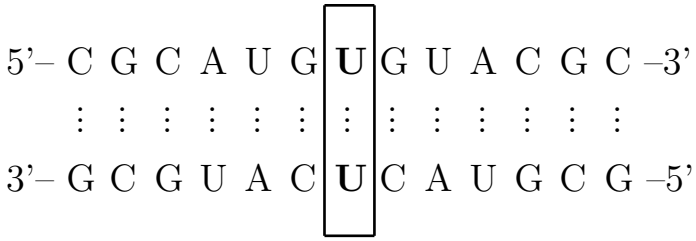


Figure 7.37: Double stranded oligonucleotide A4B4.

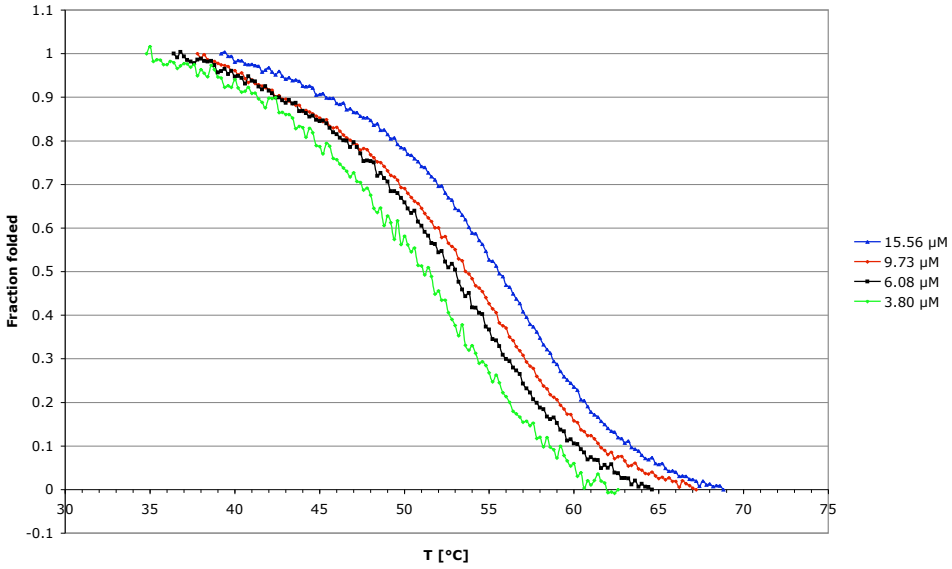


Figure 7.38: Fraction folded plot of A4B4.

Table 7.19: T_m and thermodynamic data of A4B4.

Concentration [μM]	15.56	9.73	6.08	3.80
T_m [°C]	55.6	53.8	52.6	51.6

	$\Delta G°$[kcal/mol]	$\Delta H°$[kcal/mol]	$\Delta S°$[cal/mol]	R^2
ln C_{Total} vs. $1/T_m$	-14.5	-76.4	-207	0.982
$1/T$ vs. ln K_a	-14.6 ± 0.2	-77.9 ± 2.0	-212 ± 6	

7.20 Base pairing of A4B5

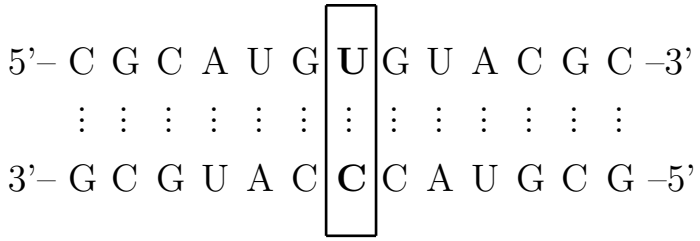


Figure 7.39: Double stranded oligonucleotide A4B5.

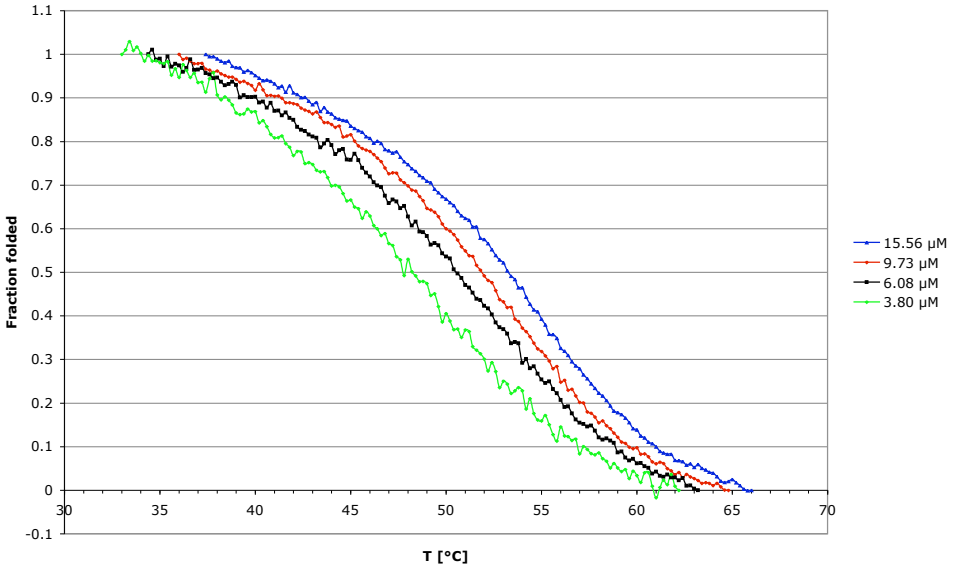


Figure 7.40: Fraction folded plot of A4B5.

Table 7.20: T_m and thermodynamic data of A4B5.

Concentration [μM]	15.56	9.73	6.08	3.80
T_m [°C]	53.2	51.8	50.6	49.0

	$\Delta G°$[kcal/mol]	$\Delta H°$[kcal/mol]	$\Delta S°$[cal/mol]	R^2
ln C_{Total} vs. $1/T_m$	-13.5	-70.9	-193	0.997
$1/T$ vs. ln K_a	-13.5 ± 0.2	-70.5 ± 2.3	-191 ± 7	

7.21 Base pairing of A5B1

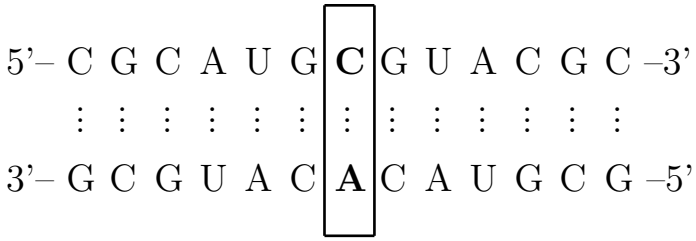


Figure 7.41: Double stranded oligonucleotide A5B1.

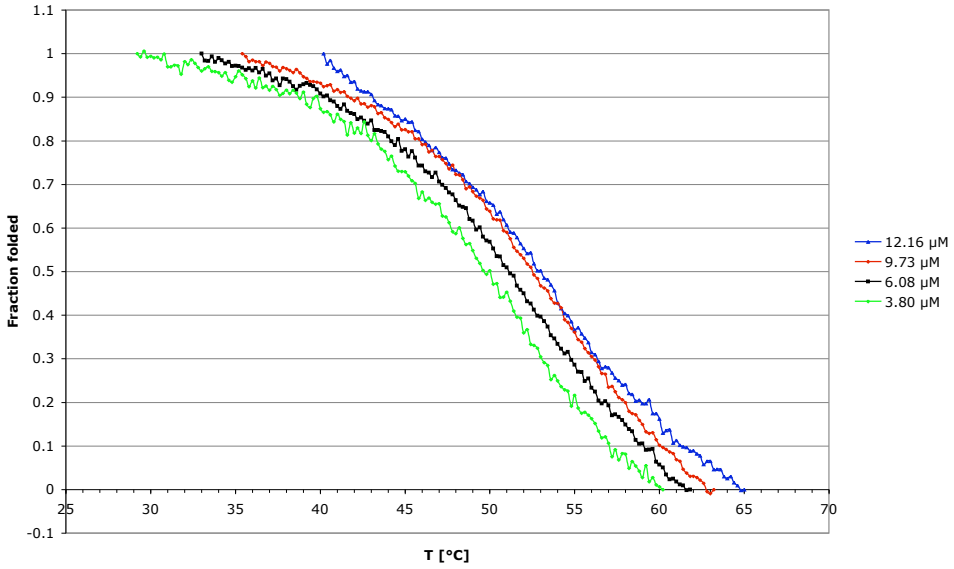


Figure 7.42: Fraction folded plot of A5B1.

Table 7.21: T_m and thermodynamic data of A5B1.

Concentration [μM]	12.16	9.73	6.08	3.80
T_m [°C]	53.0	52.5	51.1	49.6

	$\Delta G°$ [kcal/mol]	$\Delta H°$ [kcal/mol]	$\Delta S°$ [cal/mol]	R^2
ln C_{Total} vs. $1/T_m$	-13.6	-70.9	-192	0.997
$1/T$ vs. ln K_a	-13.6 ± 0.1	-70.8 ± 1.0	-192 ± 3	

7.22 Base pairing of A5B2

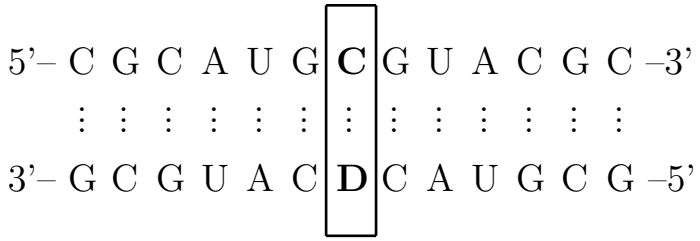


Figure 7.43: Double stranded oligonucleotide A5B2.

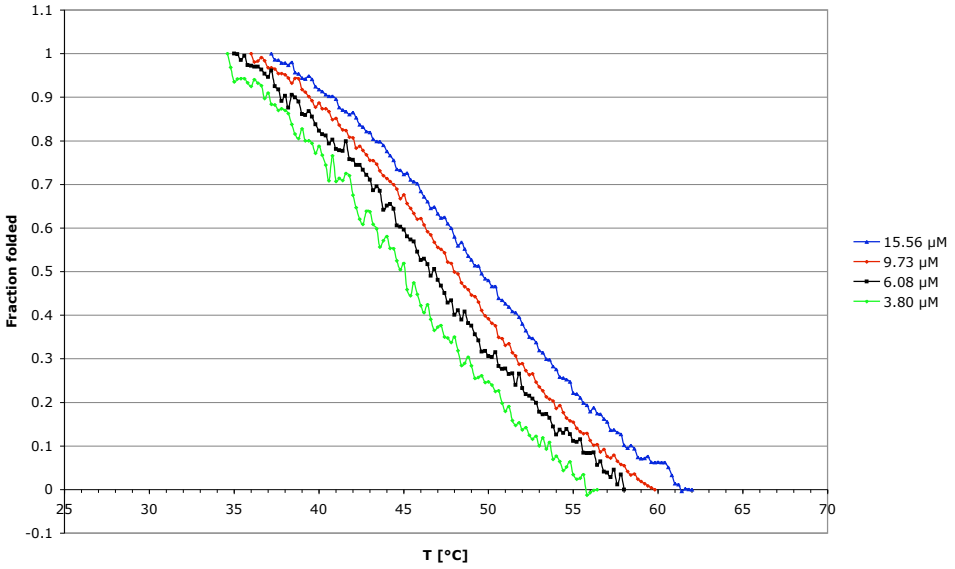


Figure 7.44: Fraction folded plot of A5B2.

Table 7.22: T_m and thermodynamic data of A5B2.

Concentration [μM]	15.56	9.73	6.08	3.80
T_m [°C]	49.3	48.0	46.6	45.1

	$\Delta G°$ [kcal/mol]	$\Delta H°$ [kcal/mol]	$\Delta S°$ [cal/mol]	R^2
ln C_{Total} vs. $1/T_m$	-12.6	-68.5	-188	0.999
$1/T$ vs. ln K_a	-12.8 ± 0.1	-70.3 ± 1.5	-193 ± 5	

7.23 Base pairing of A5B3

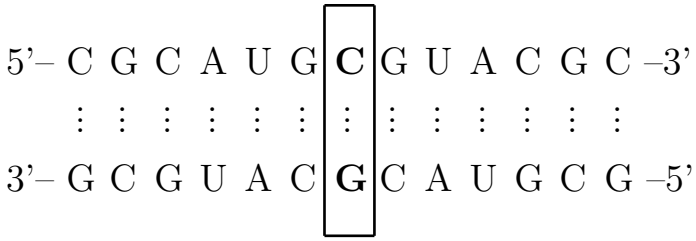


Figure 7.45: Double stranded oligonucleotide A5B3.

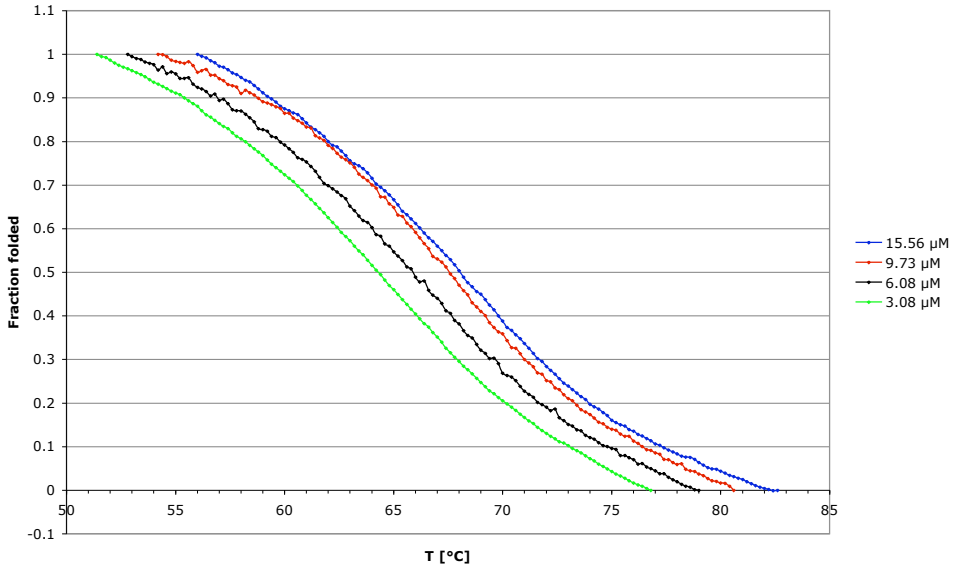


Figure 7.46: Fraction folded plot of A5B3.

Table 7.23: T_m and thermodynamic data of A5B3.

Concentration [μM]	15.56	9.73	6.08	3.80
T_m [°C]	68.4	67.6	65.8	64.4

	$\Delta G°[kcal/mol]$	$\Delta H°[kcal/mol]$	$\Delta S°[cal/mol]$	R^2
ln C_{Total} vs. $1/T_m$	-17.6	-79.4	-207	0.980
$1/T$ vs. ln K_a	-17.5 ± 0.1	-79.4 ± 0.8	-208 ± 2	

7.24 Base pairing of A5B4

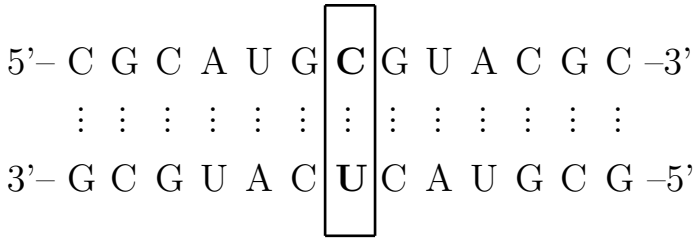


Figure 7.47: Double stranded oligonucleotide A5B4.

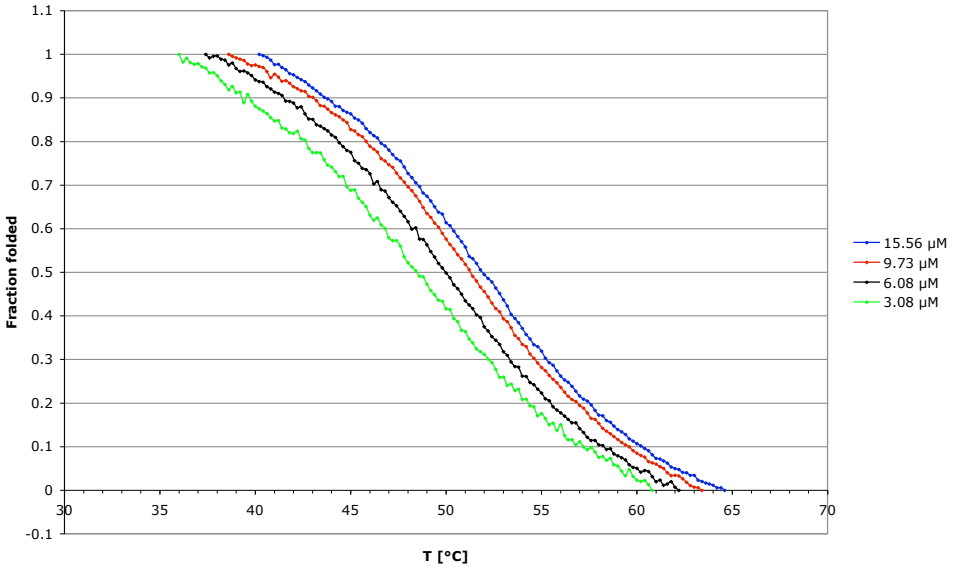


Figure 7.48: Fraction folded plot of A5B4.

Table 7.24: T_m and thermodynamic data of A5B4.

Concentration [μM]	15.56	9.73	6.08	3.80
T_m [°C]	52.2	51.4	50.0	48.6

	$\Delta G°$[kcal/mol]	$\Delta H°$[kcal/mol]	$\Delta S°$[cal/mol]	R^2
ln C_{Total} vs. $1/T_m$	-14.0	-79.4	-219	0.985
$1/T$ vs. ln K_a	-14.0 ± 0.1	-78.5 ± 1.6	-216 ± 5	

7.25 Base pairing of A5B5

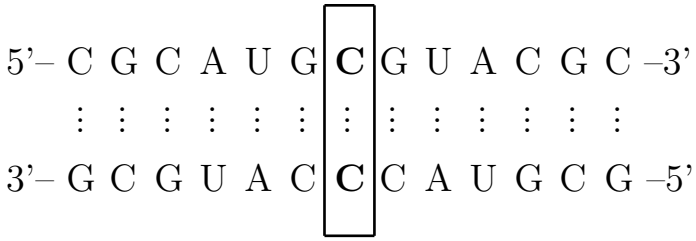


Figure 7.49: Double stranded oligonucleotide A5B5.

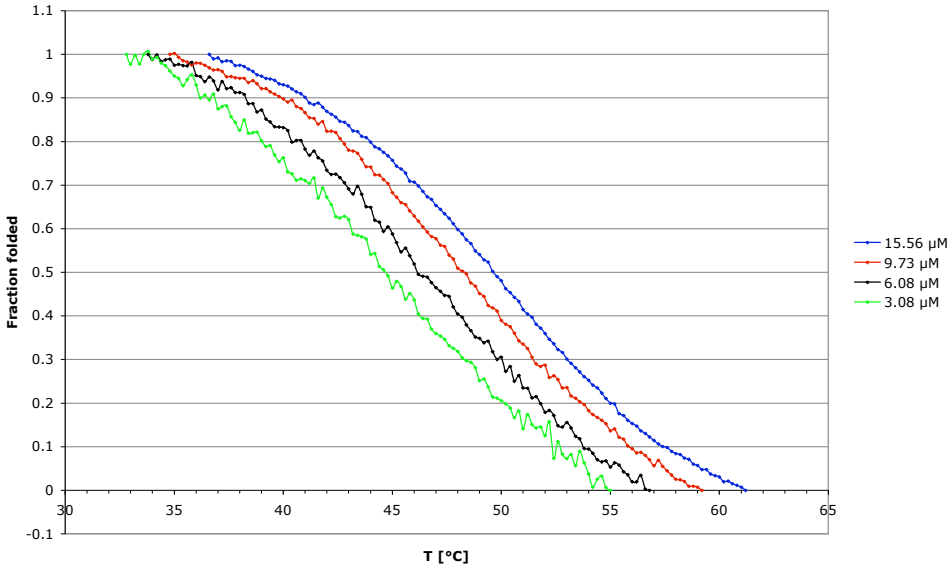


Figure 7.50: Fraction folded plot of A5B5.

Table 7.25: T_m and thermodynamic data of A5B5.

Concentration [μM]	15.56	9.73	6.08	3.80
T_m [°C]	49.4	48.0	46.8	45.4

	$\Delta G°$[kcal/mol]	$\Delta H°$[kcal/mol]	$\Delta S°$[cal/mol]	R^2
ln C_{Total} vs. $1/T_m$	-13.0	-73.5	-203	0.999
$1/T$ vs. ln K_a	-13.2 ± 0.2	-76.8 ± 1.0	-213 ± 3	

Chapter 8

Crystallographic Data of 13α

Crystallised from	CH_2Cl_2 / MeOH
Empirical formula	$C_{24}H_{27}N_3O_3$
Formula weight [g mol^{-1}]	405.49
Crystal colour, habit	colourless, prism
Crystal dimensions [mm]	0.17 x 0.30 x 0.32
Temperature [K]	160(1)
Crystal system	orthorhombic
Space group	$P2_12_12_1$ (#19)
Z	4
Reflections for cell determination	3568
2θ range for cell determination [°]	4–60
Unit cell parameters a [Å]	7.3020(1)
b [Å]	12.6767(2)
c [Å]	23.3601(4)
α [°]	90
β [°]	90
γ [°]	90
V [Å^3]	2162.33(6)
$F(000)$	864
D_X [g cm^{-1}]	1.245
μ(Mo $K\alpha$) [mm^{-1}]	0.0830
Scan type	ϖ

$2\theta_{(max)}$ [°]	60
Total reflections measured	29521
Symmetry independent reflections	3582
R_{int}	0.068
Reflections with $I < 2\sigma(I)$	3233
Reflections used in refinement	3582
Parameters refined	280
Final $R(F)$ [$I < 2\sigma(I)$ reflections]	0.0639
$wR(F^2)$ (all data)	0.1712
Weigths: $w = [\sigma^2({F_o}^2) + (0.1111P)^2 + 0.5951P]^{-1}$	where $P = ({F_o}^2 + 2{F_c}^2)/3$
Goodness of fit	0.101
Final Δ_{max}/σ	0.001
$\Delta\rho$ (max; min) [e Å^{-3}]	0.38; -0.35
$\sigma(d_{C-C}$ [Å]	0.003–0.006

Appendix

List of Figures

List of Tables

References

1. Hysell, M.; Siegel, J. S.; Tor, Y. *Org. Biomol. Chem.* **2005**, *3*, 2946–2952.

2. Knoll, A. H.; Barghoorn, E. S. *Science* **1977**, *198*, 596.

3. Haldane, J. B. S. *The Causes of Evolution*; Longmans, Green & Co, 1932.

4. Oparin, A. *The Origin of Life*; Macmillan, London, 1938.

5. Bada, J. L.; Lazcano, A. *Science* **2002**, *296*, 1982–1983.

6. Stetter, K. O. *Evolution of Hydrothermal Ecosystems on Earth (and Mars?)*; Wiley, Chichester, U.K., 1996.

7. Simoneit, B. R. T. *Adv. Space Res.* **2004**, *33*, 88–94.

8. Ferris, J. P.; Hill, A. R.; Liu, R.; Orgel, L. E. *Nature* **1996**, *381*, 59–61.

9. Cairns-Smith, A. G. *J. Theor. Biol.* **1966**, *10*, 53–88.

10. Luisi, P. L. *Orig. Life Evol. Biosphere* **1996**, *26*, 272–273.

11. Ricardo, A.; Carrigan, M. A.; Olcott, A. N.; Benner, S. A. *Science* **2004**, *303*, 196.

12. Wächtershäuser, G. *Microbiol. Rev.* **1988**, *52*, 452–484.

13. Hagmann, M. *Science* **2002**, *295*, 2006–2007.

14. Wächtershäuser, G. *Science* **2000**, *289*, 1307–1308.

15. Miller, S. L. *Science* **1953**, *117*, 528–529.

16. Miller, S. L.; Urey, H. C. *Science* **1959**, *130*, 245–251.

17. Blackmond, D. G. *Proc. Natl. Acad. Sci.* **2004**, *101*, 5732–5736.

18. Saghatelian, A.; Yokobayashi, Y.; Soltani, K.; Ghadiri, M. R. *Nature* **2001**, *409*, 797–801.

19. Pizzarello, S.; Weber, A. L. *Science* **2004**, *303*, 1151.

20. Siegel, J. S. *Nature* **2001**, *409*, 777–778.

21. Siegel, J. S. *Chirality* **1998**, *10*, 24–27.

22. Mason, S. F. *Chemical Evolution*; Clarendon Press, Oxford, 1992.

23. Guerrier-Takada, C.; Gardiner, K.; Marsh, T.; Pace, N.; Altman, S. *Cell* **1983**, *35*, 8498–57.

24. Guerrier-Takada, C.; Altman, S. *Science* **1984**, *223*, 285–286.

25. Westheimer, F. H. *Nature* **1986**, *319*, 534–536.

26. Kruger, K.; Grabowski, P. J.; Zaug, A. J.; Sands, J.; Gottschling, D. E.; Cech, T. R. *Cell* **1982**, *31*, 147–157.

27. Cech, T. R. *Int. Rev. Cytol.* **1985**, *93*, 3–22.

28. Zaug, A. J.; Cech, T. R. *Science* **1986**, *231*, 470–475.

29. Gilbert, W. *Nature* **1986**, *319*, 618.

30. Sharp, P. A. *Cell* **1985**, *42*, 397–400.

31. White, H. B. *J. Mol. Evol.* **1976**, *7*, 101–104.

32. Pääbo, S. *Scientific American* **1993**, *269*, 86–93.

33. Schwartz, A. W. *The Molecular Origins of Life*; Cambridge University Press, 1998.

34. Nielson, P. E. *Origins Life Evol. Biosphere* **1993**, *23*, 323–327.

35. Nelson, K. E.; Levy, M.; Miller, S. L. *Proc. Natl. Acad. Sci.* **2000**, *97*, 3868–3871.

36. Kim, S. K.; Nielson, P. E.; Egholm, M.; Buchardt, O.; Berg, R. H.; Norden, B. *J. Am. Chem. Soc.* **1993**, *115*, 6477–6481.

37. Nielson, P. E.; Egholm, M.; Berg, R. H.; Buchardt, O. *Science* **1991**, *254*, 1497–1500.

38. Ueda, N.; Kawabata, T.; Takemoto, K. *J. Heterocycl. Chem.* **1971**, *8*, 827–829.

39. Schwartz, A. W.; Orgel, L. E. *Science* **1985**, *228*, 585–587.

40. Seita, T.; Yamauchi, K.; Kinoshita, M.; Imoto, M. *Macromol. Chem.* **1972**, *154*, 255–261.

41. Zhang, L.; Peritz, A.; Meggers, E. *J. Am. Chem. Soc.* **2005**, *127*, 4174–4175.

42. Westheimer, F. H. *Science* **1987**, *235*, 1173–1178.

43. Usher, D. A. *Nature New Biol.* **1972**, *235*, 207–208.

44. Usher, D. A. *Science* **1977**, *196*, 311–313.

45. Eschenmoser, A. *Origins Life Evol. Biosphere* **2004**, *34*, 277–306.

46. Eschenmoser, A.; Dobler, M. *Helv. Chim. Acta* **1992**, *75*, 218–259.

47. Hunziker, J.; Roth, H.-J.; Böhringer, M.; Giger, A.; Diederichsen, U.; Göbel, M.; Krishnan, R.; Jaun, B.; Leumann, C.; Eschenmoser, A. *Helv. Chim. Acta* **1993**, *76*, year.

48. Böhringer, M.; Roth, H.-J.; Hunziker, J.; Göbel, M.; Krishnan, R.; Giger, A.; Schweizer, B.; Schreiber, J.; Leumann, C.; Eschenmoser, A. *Helv. Chim. Acta* **1992**, *75*, 1416–1477.

49. Eschenmoser, A.; Loewenthal, E. *Chem. Soc. Rev.* **1992**, *21*, 1–16.

50. Schönig, K.-U.; Scholz, P.; Guntha, S.; Wu, X.; Krishnamurthy, R.; Eschenmoser, A. *Science* **2000**, *290*, 1347–1351.

51. Eschenmoser, A. *Science* **1999**, *284*, 2118–2124.

52. VanAerschot, A.; Verheggen, I.; Hendrix, C.; Herdewijn, P. *Angew. Chem. Int. Ed.* **1995**, *34*, 1338–1339.

53. Orgel, L. E. *Science* **2000**, *290*, 1306–1307.

54. Piccirilli, J. A.; Krauch, T.; Moroney, S. E.; Benner, S. A. *Nature* **1990**, *343*, 33–37.

55. Kolb, V. M.; Dworkin, J. P.; Miller, S. L. *J. Mol. Evol.* **1994**, *38*, 549–557.

56. Wächtershäuser, G. *Proc. Natl. Acad. Sci.* **1988**, *85*, 1134–1135.

57. Inoue, T.; Orgel, L. E. *Science* **1983**, *219*, 859–862.

58. Hatfield, D.; Wyngaarden, J. B. *J. Biol. Chem.* **1964**, *239*, 2580–2592.

59. Zimmerman, M. *J. Biol. Chem.* **1966**, *241*, 4914–4916.

60. Haraguchi, K.; Greenberg, M. M. *J. Am. Chem. Soc.* **2001**, *123*, 8636–8637.

61. Kubo, I.; Kim, M.; Wood, W. F.; Naoki, H. *Tetrahedron Lett.* **1986**, *27*, 4277–4280.

62. Löpfe, M. Ph.D. thesis, University of Zurich, 2008.

63. Hysell, M.; Siegel, J. S.; Tor, Y. *Org. Biomol. Chem.* **2005**, *3*, 2946–2952.

64. Greenberg, M. M.; Hantosi, Z.; Wiederholt, C. J.; Rithner, C. D. *Biochemistry* **2001**, *40*, 15856–15861.

65. Delaney, M. O.; Wiederholt, C. J.; Greenberg, M. M. *Angew. Chem. Int. Ed.* **2002**, *41*, 771–773.

66. Berstis, L. M.Sc. thesis, University of Zurich, 2009.

67. Darwin, C. R. *On the Origin of the Species by Means of Natural Selection, or, the Preservation of Favoured Races in the Struggle for Life*; John Murray, London, 1859.

68. Mendel, G. *Verh. Nat.forsch. Ver. Brünn* **1866**, *4*, 3–47.

69. Haeckel, E. *Generelle Morphologie der Organismen*; Reimer, Berlin, 1866.

70. His, W. *Die Histochemischen und Physiologischen Arbeiten von Friedrich Miescher - Aus dem wissenschaftlichen Briefwechsel von F. Miescher*; F. C. W. Vogel, Leipzig, 1897; Vol. 1.

71. Miescher, F. *Med.-Chem. Unters.* **1971**, *4*, 441–460.

72. Dahm, R. *Developmental Biology* **2005**, *278*, 274–288.

73. Flemming, W. *Arch. Pathol. Anat. Physiol.* **1879**, *77*, 1–29.

74. Boveri, T. *Jena. Zeit. Naturwiss.* **1888**, *22*, 685–882.

75. Kay, L. E. *The Molecular Vision of Life*; Oxford University Press, 1992.

76. Griffith, F. *J. Hyg.* **1928**, *27*, 113.

77. Avery, O. T.; MacLeod, C. M.; McCarty, M. *J. Exp. Med.* **1944**, *79*, 137–159.

78. Vischer, E.; Zamenhof, S.; Chargaff, E. *J. Biol. Chem.* **1949**, *177*, 429.

79. Palmiter, R. D.; Brinster, R. L.; Hammer, R. E.; Trumbauer, M. E.; Rosenfeld, M. G.; Birnberg, N. C.; Evans, R. M. *Nature* **1982**, *300*, 611–615.

80. Franklin, R. E.; Gosling, R. G. *Nature* **1953**, *171*, 740–741.

81. Watson, J. D.; Crick, F. H. C. *Nature* **1953**, *171*, 737–738.

82. Wilkins, M. H.; Stokes, A. R.; Wilson, H. R. *Nature* **1953**, *171*, 738–740.

83. Crick, F. H. C. *Symp. Soc. Exp. Biol. XII* **1958**, 139–163.

84. Maxam, A. M.; Gilbert, W. *Proc. Natl. Acad. Sci.* **1977**, *74*, 560–564.

85. Sanger, F.; Nicklen, S.; Coulson, A. R. *Proc. Natl. Acad. Sci.* **1977**, *74*, 5463–5467.

86. *Science* **2001**, *291*, 1218.

87. Voet, D.; Voet, J. G. *Biochemie*; John Wiley & Sons, New York, 2004; Vol. 1. korrigierter Nachdruck der 1. Auflage.

88. Langen, P.; Hucho, F. *Angew. Chem. Int. Ed.* **2008**, *47*, 1824–1827.

89. Gordon, J. L. *Biochem. J.* **1986**, *233*, 309–319.

90. Baltimore, D. *Nature* **1970**, *226*, 1209–1211.

91. Temin, H. M.; Mizutani, S. *Nature* **1970**, *226*, 1211–1213.

92. *Biochemistry* **1970**, *9*, 4022–4027.

93. Sierzputowska-Gracz, H.; Gopal, H. D.; Agris, P. F. *Nucleic Acids Res.* **1986**, *14*, year.

94. Sober, H. A.; Harte, R. A.; Sober, E. K. *Handbook of Biochemistry. Selected Data for Molecular Biology*; Chem. Rubber Co., Cleveland, Ohio, 1970.

95. Izatt, R. M.; Christensen, J. J.; Rytting, J. H. *Chem. Rev.* **1971**, *71*, 439–481.

96. Nakajima, T.; Pullman, B. *J. Am. Chem. Soc.* **1959**, *81*, 3876–3877.

97. Cochran, W. *Acta Cryst.* **1951**, *4*, 81.

98. Jardetzky, C. D.; Jardetzky, O. *J. Am. Chem. Soc.* **1960**, *82*, 222–229.

99. Velikyan, I.; Acharya, S.; Trifonova, A.; Foeldesi, A.; Chattopadhyaya, J. *J. Am. Chem. Soc.* **2001**, *123*, 2893–2894.

100. Elguero, J.; Marzin, C.; Katritzky, A. R.; Linda, P. *The Tautomerism of Heterocycles*; Academic Press, New York, 1976.

101. Chon, S. H.; Reid, B. R.; Brown, T.; Hunter, W. N. *Oxford Handbook of Nucleic Acid Structure*; Oxford Science Publishing, Oxford, 1999.

102. Topal, M. D.; Fresco, J. R. *Nature* **1976**, *263*, 285–289.

103. Alyoubi, A. O.; Hilal, R. H. *Biophys. Chem.* **1995**, *55*, 231–237.

104. Sundaralingam, M. *Biopolymers* **1969**, *7*, 821–860.

105. Saenger, W. *Principles of Nucleic Acid Structure*; Springer Verlag, New York, 1984.

106. Klein, W.; Thannhauser, S. J. *Z. physiol. Chem.* **1933**, *218*, 173–180.

107. Klein, W.; Thannhauser, S. J. *Z. physiol. Chem.* **1934**, *224*, 252–260.

108. Franklin, R. E.; Gosling, R. G. *Nature* **1953**, *172*, 156–157.

109. Wang, A. H.-J.; Quigley, G. J.; Kolpak, F. J.; Crawford, J. L.; van Boom, J. H.; van der Marel, G.; Rich, A. *Nature* **1979**, *282*, 680–686.

110. Wing, R.; Drew, H.; Takano, T.; Broka, C.; Tanaka, S.; Itakura, K.; Dickerson, R. E. *Nature* **1980**, *287*, 755–758.

111. Hobza, P.; Sponer, J. *Chem. Rev.* **1999**, *99*, 3247–3276.

112. Petersheim, M.; Turner, D. H. *Biochemistry* **1983**, *22*, 256–263.

113. Chalikian, T. V.; Volker, J.; Plum, G. E.; Breslauer, K. J. *Proc. Natl. Acad. Sci.* **1999**, *96*, 7853–7858.

114. Stewart, R. F.; Jensen, L. H. *J. Chem. Phys.* **1964**, *40*, 2071–2075.

115. Kool, E. T. *Annu. Rev. Biophys. Biomol. Struct.* **2001**, *30*, 1–22.

116. Guckian, K. M.; Ren, R. X. F.; Chaudhuri, N. C.; Tahmassebi, D. C.; Kool, E. T. *J. Am. Chem. Soc.* **2000**, *122*, 2213–2227.

117. Bubienko, E.; Cruz, P.; Thomason, J. F.; Borer, P. N. *Prog. Nucleic Acid Res. Mol. Biol.* **1983**, *30*, 41–90.

118. Yakuvchuk, P.; Protozanova, E.; Frank-Kamenetskii, M. D. *Nucleic Acids Res.* **2006**, *34*, 564–574.

119. Owczarzy, R.; You, Y.; Moreira, B. G.; Manthey, J. A.; .Huang, L.; Behlke, M. A.; Walder, J. A. *Biochemistry* **2004**, *43*, 3537–3554.

120. Schildkraut, C.; Lifson, S. *Biopolymers* **1965**, *3*, 195–208.

121. Hoogsteen, K. *Acta Cryst.* **1963**, *16*, 907–916.

122. Johnson, J. E.; Smith, J. S.; Kozak, M. L.; Johnson, F. B. *Biochimie* **2008**, *90*, 1250–1263.

123. Zagryadskaya, E. I.; Doyon, F. R.; Steinberg, S. V. *Nucleic Acids Res.* **2003**, *31*, 3946–3953.

124. Crick, F. H. C. *J. Mol. Biol.* **1966**, *19*, 548–444.

125. Varani, G.; McClain, W. *EMBO Rep.* **2000**, *1*, 18–23.

126. Pace, N. R. *Cell* **1991**, *65*, 531–533.

127. Russell, M. J.; Daniel, R. M.; Hall, A. J.; Sherringham, J. A. *J. Mol. Evol.* **1994**, *39*, 231–243.

128. Russell, M. J.; Hall, A. J.; Cairns-Smith, A. G.; Braterman, P. S. *Nature* **1988**, *336*, 117.

129. Levy, M. L.; Miller, S. L. *Proc. Natl. Acad. Sci.* **1998**, *95*, 7933–7938.

130. Lindahl, T. *Nature* **1993**, *362*, 709–715.

131. Karran, P.; Lindahl, T. *Biochemistry* **1980**, *19*, 6005–6011.

132. Frederico, L. A.; Kunkel, T. A.; Shaw, B. R. *Biochemistry* **1990**, *29*, 2532–2537.

133. Jones, M.; Wagner, R.; Radman, M. *J. Mol. Biol.* **1987**, *194*, 155–159.

134. Lutz, W. K. *Mutat. Res.* **1990**, *238*, 287–295.

135. Lindahl, T.; Nyberg, B. *Biochemistry* **1972**, *11*, 3610–3618.

136. Lindahl, T.; Karlström, O. *Biochemistry* **1973**, *12*, 5151–5154.

137. Lindahl, T. *J. Biol. Chem.* **1967**, *242*, 1970–1973.

138. Gasparutto, D.; Ravanat, J.-L.; Gérot, O.; Cadet, J. *J. Am. Chem. Soc.* **1998**, *120*, 10283–10286.

139. Townsend, L. B.; Robins, R. K. *J. Am. Chem. Soc.* **1963**, *85*, 242–243.

140. Ghose, A. K.; Viswanadhan, V. N.; Sanghvi, Y. S.; Nord, L. D.; Willis, R. C.; Revankar, G. R.; Robins, R. K. *Proc. Natl. Acad. Sci.* **1989**, *86*, 8242.

141. Berman, H. M.; Rousseau, R. J.; Mancuso, R. W.; Kreishman, G. P.; Robins, R. K. *Tetrahedron Lett.* **1973**, *14*, 3099.

142. Gajewski, E.; Rao, G.; Nackerdien, Z.; Dizdaroglu, M. *Biochemistry* **1990**, *29*, 7876–7882.

143. Boiteux, S.; Gajewski, E.; Laval, J.; Dizdaroglu, M. *Biochemistry* **1992**, *31*, 106–110.

144. Doetsch, P. W.; Zastawny, T. H.; Martin, A. M.; Dizdaroglu, M. *Biochemistry* **1995**, *34*, 737–742.

145. Krakaya, A.; Jaruga, P.; Bohr, V. A.; Grollman, A. P.; Dizdaroglu, M. *Nucleic Acids Res.* **1997**, *25*, 474–479.

146. Graziewicz, M. A.; Zastawny, T. H.; Olinski, R.; Tudek, B. *Mutat. Res. - DNA Repair* **1999**, *434*, 41–52.

147. Kalam, M. A.; Haraguchi, K.; Chandani, S.; Loechler, E. L.; Moriya, M.; Greenberg, M. M.; Basu, A. K. *Nucleic Acids Res.* **2006**, *34*, 2305–2315.

148. Shin, H.; Min, C. *Modified Nucleosides in Biochemistry, Biotechnology and Medicine*; Wiley-VCH, 2008.

149. Fortin, H.; Tomasi, S.; Delcros, J.-G.; Bansard, J.-Y.; Boustie, J. *ChemMedChem* **2006**, *1*, 189–196.

150. Robertson, M. P.; Levy, M. L.; Miller, S. L. *J. Mol. Evol.* **1996**, *43*, 543–550.

151. Shapiro, R. *Proc. Natl. Acad. Sci.* **1999**, *96*, 4396–4401.

152. Shapiro, R.; Servis, R. E.; Welcher, M. *J. Am. Chem. Soc.* **1970**, *92*, 422–424.

153. Burgdorf, L. T.; Carell, T. *Chem. Eur. J.* **2002**, *8*, 293–301.

154. Raoul, S.; Bardet, M.; Cadet, J. *Chem. Res. Toxicol.* **1995**, *7*, 924–933.

155. Berger, M.; Cadet, J. *Z. Naturforsch., B: Chem. Sci.* **1985**, *40*, 1519–1531.

156. *Carbohydrate Res.* **1987**, *171*, year.

157. Wellington, K. W.; Benner, S. A. *Nucleosides, Nucleotides, and Nucleic Acids* **2006**, *25*, 1309–1333.

158. Townsend, L. B. *Chemistry of Nucleosides and Nucleotides*; Plenum Press, New York, 1994.

159. Aketani, S.; Tanaka, K.; Yamamoto, K.; Ishihama, A.; Cao, H.; Tengeiji, A.; Hiraoka, S.; Shiro, M.; Shionoya, M. *J. Med. Chem.* **2002**, *45*, 5594–5603.

160. Doi, Y.; Chiba, J.; Morikawa, T.; Inouye, M. *J. Am. Chem. Soc.* **2008**, *130*, 8762–8768.

161. Morales, J. C.; Kool, E. T. *J. Am. Chem. Soc.* **2000**, *122*, 1001–1007.

162. Wichai, U.; Woski, S. A. *Org. Lett.* **1999**, *1*, 1173.

163. Gensler, W. J.; Chan, S.; Ball, D. B. *J. Am. Chem. Soc.* **1975**, *97*, 436–437.

164. Gensler, W. J.; Chan, S.; Ball, D. B. *J. Org. Chem.* **1981**, *46*, 3407.

165. Sato, T.; Noyori, R. *Tetrahedron Lett.* **1979**, *20*, 3996.

166. Secrist, J. A. *J. Org. Chem.* **1978**, *43*, 2925.

167. Cupps, T. L.; Wise, D. S.; Townsend, L. B. *J. Org. Chem.* **1982**, *47*, 5115–5120.

168. Cupps, T. L.; Wise, D. S.; Townsend, L. B. *J. Org. Chem.* **1986**, *51*, 1058–1064.

169. Haraguchi, K.; Delaney, M. O.; Wiederholt, C. J.; Sambandam, A.; Hantosi, Z.; Greenberg, M. M. *J. Am. Chem. Soc.* **2002**, *124*, 3263–3269.

170. Delaney, M. O.; Greenberg, M. M. *Chem. Res. Toxicol.* **2002**, *15*, 1460–1465.

171. Löpfe, M.; Bischof, B.; Siegel, J. S. *Org. Biomol. Chem.* **submitted**.

172. Michelson, A. M.; Todd, A. R. *J. Chem. Soc.* **1955**, 2632–2638.

173. Letsinger, R. L.; Finnan, J. L.; Heavner, G. A.; Lunsford, W. B. *J. Am. Chem. Soc.* **1975**, *97*, 3278–3279.

174. Letsinger, R. L.; Lunsford, W. B. *J. Am. Chem. Soc.* **1976**, *98*, 3655–3661.

175. Beaucage, S. L.; Caruthers, M. H. *Tetrahedron Lett.* **1981**, *22*, 1859–1862.

176. Beaucage, S. L.; Iyer, R. *Tetrahedron* **1992**, *48*, 2223–2311.

177. Bredereck, H.; Effenberger, F.; Simchen, G. *Chem. Ber.* **1962**, *135*, 1078–1080.

178. Gold, A.; Sangaiah, R. *Nucleosides and Nucleotides* **1990**, *9*, 907–912.

179. Dyson, M. R.; Coe, P. L.; Walker, R. T. *Carbohydrate Res.* **1991**, *216*, 237–248.

180. Cazeau, P.; Duboudin, F.; Moulines, R.; Babot, O.; Dunogues, J. *Tetrahedron* **1987**, *43*, 2075–2088.

181. Hoffmann, M. G.; Schmidt, R. R. *Liebigs Ann. Chem.* **1985**, 2403.

182. Niedballa, U.; Vorbrüggen, H. *J. Org. Chem.* **1974**, *39*, 3654–3660.

183. Shao, H.; Wang, Z.; Lacroix, E.; Wu, S.-H.; Jennings, H. J.; Zou, W. *J. Am. Chem. Soc.* **2002**, *124*, 2130–2131.

184. Lorthiois, E.; Marek, I.; Normant, J. F. *J. Org. Chem.* **1998**, *63*, 2442–2450.

185. Abdulla, R. F.; Brinkmeyer, R. S. *Tetrahedron* **1979**, *35*, 1675–1735.

186. Ciuffreda, P.; Casati, S.; Manzocchi, A. *Magnetic Res. Chem.* **2007**, *45*, 781–784.

187. Gait, M. J. *Oligonucleotide Synthesis, a practical approach*; IRL Press, Oxford, 1984.

188. Loren, J. Ph.D. thesis, University of San Diego, 2003.

189. Camaioni, E.; Costanzi, S.; Vittori, S.; Volpini, R.; Klotz, K.-N.; Cristalli, G. *Bioorg. Med. Chem.* **1998**, *6*, 523–533.

190. Harada, H. et al. *J. Med. Chem.* **2001**, *44*, 170–179.

191. Harada, H.; Asano, O.; Kawata, T.; Inoue, T.; Horizoe, T.; Yasuda, N.; Nagata, K.; Murakami, M.; Nagaoka, J.; Kobayashi, S.; Tanaka, I.; Abe, S. *Bioorg. Med. Chem.* **2001**, *9*, 2709–2726.

192. Matsuda, A.; Shinozaki, M.; Yamaguchi, T.; Homma, H.; Nomoto, R.; Miyasaka, T.; Watanabe, Y.; Abiru, T. *J. Med. Chem.* **1992**, *35*, 241–252.

193. Nair, V.; Young, D. A.; DeSilva, R. *J. Org. Chem.* **1987**, *52*, 1344–1347.

194. Nair, V.; Richardson, S. G. *Synthesis* **1982**, *8*, 670–672.

195. Klein, R. S.; Fox, J. J. *J. Org. Chem.* **1972**, *37*, 4381–4386.

196. Bredereck, H.; Gompper, R.; Herlinger, H. *Chem. Ber.* **1958**, *91*, 2832–2849.

197. Ogata, M.; Watanabe, H.; Tori, K.; Kano, H. *Tetrahedron Lett.* **1964**, 19–24.

198. Yamanaka, H.; Ogawa, S.; Sakamoto, T. *Heterocycles* **1981**, *16*, 573.

199. Yamanaka, H.; Sakamoto, T.; Niitsoma, S. *Heterocycles* **1990**, *31*, 923.

200. Czernecki, S.; Ville, G. *J. Org. Chem.* **1989**, *54*, 610.

201. Song, J.; Hollingsworth, R. I. *Tetrahedron Asym.* **2001**, *12*, 387.

202. Walker, J. A.; Chen, J. J.; Wise, D. S.; Townsend, L. B. *J. Org. Chem.* **1996**, *61*, 2219.

203. Murray, T. P.; Hay, J. V.; Portlock, D. E.; Wolfe, J. F. *J. Org. Chem.* **1974**, *39*, 595–600.

204. Hainke, S.; Arndt, S.; Seitz, O. *Org. Biomol. Chem.* **2205**, *3*, 4233–4238.

205. Vesnaver, G.; Chang, C.-N.; Eisenberg, M.; Grollman, A. P.; Breslauer, K. J. *Biochemistry* **1989**, *86*, 3614–3618.

206. Vesnaver, G.; Breslauer, K. J. *Biochemistry* **1991**, *88*, 3569–3573.

207. Plum, G. E.; Grollman, A. P.; Johnson, F.; Breslauer, K. J. *Biochemistry* **1995**, *34*, 16148–16160.

208. Thomas, R. *Experientia* **1951**, *7*, 261–263.

209. Nesbet, R. K. *Mol. Phys.* **1964**, *7*, 211–221.

210. Mergny, J.-L.; Lacroix, L. *Oligonucleotides* **2003**, *13*, 515–537.

211. Mergny, J.-L.; Maurizot, J.-C. *Chem. Bio. Chem.* **2001**, *2*, 124–132.

212. Cahen, P.; Luhmer, M.; Fontaine, C.; Morat, C.; Reisse, J.; Bartik, K. *Biophys. J.* **2000**, *78*, 1059–1069.

213. Guo, Q.; Lu, M.; Kallenbach, N. R. *J. Biol. Chem.* **1992**, *267*, 15293–15300.

214. Duguid, J. G.; Bloomfield, V. A.; Benevides, J. M.; Thomas, G. J. *Biophys. J.* **1996**, *71*, 3350–3360.

215. Manning, G. S. *Biopolymers* **1972**, *11*, 937–949.

216. DePrisco Albergo, D.; Marky, L. A.; Breslauer, K. J.; Turner, D. H. *Biochemistry* **1981**, *20*, 1409–1413.

217. Marky, L. A.; Breslauer, K. J. *Biopolymers* **1987**, *26*, 1601–1620.

218. Chaires, J. B. *Biophys. Chem.* **1997**, *64*, 15–23.

219. Vesnaver, G.; Breslauer, K. J. *Proc. Natl. Acad. Sci.* **1991**, *88*, 3569–3573.

220. Bhat, C. C. *Synthetic Procedures in Nucleic Acid Chemistry*; John Wiley & Sons, New York, 1968; Vol. 1, pp. 521-522.

221. Czernecki, S.; Georgoulis, C.; Provelenghiou, C. *Tetrahedron Lett.* **1976**, *39*, 3535–3536.

222. Wierenga, W.; Skulnick, H. I. *Carbohydrate Res.* **1981**, *90*, 41–52.

223. Hart, D. J.; Hong, W.-P.; Hsu, L.-Y. *J. Org. Chem.* **1987**, *52*, 4665–4673.

Acknowledgment

At this point, I would like to heartily thank all the people who helped me during my PhD thesis and who made a decisive contribution to this work. In particular:

My highly esteemed doctoral adviser and mentor Prof. Dr. Jay S. Siegel for the admittance in his scientific group, the showed confidence and the excellent support.

Prof. Dr. Nathan Luedtke for his great and useful help as a DNA expert.

Prof. Dr. John A. Robinson and Prof. Dr. Oliver Zerbe for being in my Promotionskomitee.

Prof. Dr. Peter Rüedi for his help as a HPLC expert, as well as for inspiring discussions about chemistry and more important topics.

Dr. Michael Löpfe for the fantastic teamwork and the emerged friendship.

Derik Frantz for the essential proofreading.

Dr. Christian Clerc, Dr. Markus Stöckli, Peter Übelhart, Michael Wächter, Fitore Kasumaj, Roman Maag, Fabian Konrad, Frank Rogano, Kasper Moth-Poulsen and all Siegel-/Finney-group members for the enjoyable lab atmosphere and experiences in- and outside the university.

Dr. Laurent Bigler and the hole MS team for the enormous task they have done for me.

Nadja Bross and Simon Jurt for countless measured NMR spectra.

PD Dr. Anthony Linden and Sasha Blumentritt for X-ray analysis.

SNF and the Kanton Zürich for founding this project.

Regula and Christian Clerc-Brütsch and Markus Stöckli for their friendship, countless dinners, and shared diverting evenings.

My family: Monika and Kurt Bischof, Anita, Christof, and Jeannette Krauss for their encouragement, belief, and love.

Curriculum Vitae

Last name:		Bischof
First name:		Benno
Born on:		April 7, 1979 in St.Gallen
Native place:		Eggersriet/Grub SG
Nationality:		Switzerland
Civil status:		unmarried
Education:	1994 – 1998	Kantonsschule am Burggraben (St.Gallen) Matura Typus C
	1999 – 2004	Studies of chemistry, University of Zurich major subject: Organic chemistry minor subject: Inorganic chemistry second minor subject: Biochemistry
	2004	Diploma thesis in organic chemistry Prof. Dr. H. Heimgartner Organic Chemistry Institute, UZH
Diploma thesis:	Title:	"Synthese eines Aib-haltigen Endothiopeptids."
PhD thesis:	2004-2009	Prof. Dr. Jay S. Siegel, OCI, UZH

Active participation at scientific congresses and work shops:

Poster presentations:

Dorothy Crowfoot Hodgkin Symposium, May 2006, University of Zurich.

XVII International Roundtable on Nucleosides, Nucleotides and Nucleic Acids, September 2006, University of Bern.

3^{rd} Nucleic Acid Chemical Biology PhD Summer School, June 2007, University of Odense, Denmark.

Swiss Academy of Science: Rigi-Workshop: Frontiers in Chemical Biology, January 2009.

Oral presentations:

Doktorandentag of the Organic Chemistry Institute, June 2007, University of Zurich:
"The Origin of Life and Synthesis of a Homo-C Nucleotide DNA-Analogue."

SCS Fall Meeting, September 2008, University of Zurich:
"Synthesis of a Homo-C Nucleotide and Investigation of Its Base Pairing Ability."

Scientific publications:

Bärtsch, A.; Bischof, B.; Heimgartner, H. *Polish J. Chem.*, **2009**, *83*, 195–206.
"Synthesis of Endothiopeptides by Using the 'Azirine/Oxazolone Methode'."

Löpfe, M.; Bischof, B.; Siegel, J. S. *Org. Biomol. Chem.* **submitted**.

Bischof, B.; Siegel, J. S. *Org. Biomol. Chem.*, **under progress**.

Work experience:

Industrial practical in analytical chemistry, MDS Pharma Services, July – Sept. 2003, Fehraltdorf ZH, Switzerland.

Academic practical in physics about superconducting materials, with Prof. Dr. A. Schilling, University of Zurich, 2003-2004.